



Identifying the gaps and opportunities for NAMs application in regulatory toxicology

Barbara L. F. Kaplan¹, Alan Hoberman², Simone H. Stahl³,
George E. N. Kass⁴ and William Slikker Jr.⁵

The increasing number of chemicals (including those found in cosmetics and food) and pharmaceutical modalities challenges traditional animal-based approaches for hazard identification and risk assessment. New approach methodologies (NAMs) offer promising alternatives to improve efficiency, reduce animal use, and enhance human relevance in regulatory toxicology. This manuscript synthesizes insights from expert discussions and recent literature to identify gaps and opportunities for NAMs adoption. We outline the evolution of NAMs, emphasizing the importance of context of use, mechanistic understanding, and reproducibility. Current NAMs applications include skin and eye irritation, sensitization, genotoxicity, and endocrine activity, while systemic, reproductive, and developmental toxicity remain poorly addressed. Obstacles to broader implementation include technical limitations in replicating complex physiology, lack of standardized protocols, and challenges in quantitative extrapolation to human outcomes. Emerging technologies such as microphysiological systems (MPS), organ-on-chip models, and artificial intelligence (AI) frameworks show potential to bridge these gaps but require rigorous quality control and harmonized validation standards. Regulatory acceptance is hindered by inconsistent terminology, limited confidence in predictive power, and resource constraints for qualification studies. We propose a unified framework focusing on context of use, performance standards, biological variability, and reproducibility to facilitate integration of NAMs into regulatory decision-making. Collaborative efforts among developers, regulators, and stakeholders are essential to advance NAMs as reliable tools for next-generation risk assessment, ultimately supporting global initiatives to reduce animal testing while ensuring human (and animal) safety.

Addresses

¹ Center for Environmental Health Sciences, Department of Comparative Biomedical Sciences, College of Veterinary Medicine, Mississippi State University, Mississippi State, MS, USA

² Argus International, Inc. Mechanicsville, PA, USA

³ ApconiX Ltd, Alderley Park, Alderley Edge, SK10 4TG, United Kingdom

⁴ School of Medicine, Pharmacology & Therapeutics, University of Galway, Galway, Ireland and Retired from the European Food Safety Authority, Parma, Italy

⁵ Retired from the National Center for Toxicological Research/FDA, Little Rock, AR, USA

Corresponding author: Kaplan, Barbara L.F. (BKaplan@cvm.msstate.edu)

Current Opinion in Toxicology 2026, **47**:100591

This review comes from a themed issue on **VSI: Next Generation Risk Assessment**

Edited by **Emanuela Corsini, Heather Wallace, and Maria Lucia Zaidan Dagli**

Available online 20 May 2026

For a complete overview see the [Issue](#) and the [Editorial](#)

<https://doi.org/10.1016/j.cotox.2026.100591>

2468-2020/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

Keywords

NAMs, Next generation risk assessment, Regulatory toxicology.

Abbreviations

ADME, absorption, distribution, metabolism, excretion; AI, artificial intelligence; DILI, drug-induced liver injury; ECVAM, European Centre for the Validation of Alternative Methods (EU); FDCA, federal food drug and cosmetic act (US); GD, guidance document; ICH, International Council for Harmonisation; ICCVAM, Interagency Coordinating Committee on the Validation of Alternative Methods; IVIVE, *in vitro* to *in vivo* extrapolation; ML, machine learning; MPS, microphysiological systems; NAMs, new approach methodologies; OECD, Organisation for Economic Cooperation and Development; QSAR, quantitative structure activity relationships; REACH, registration, evaluation, authorisation, and restriction of chemicals (EU); TG, test guideline; TSCA, toxic substances control act (US); WoE, weight of evidence.

Introduction

Regulatory toxicology has been defined as the application of the knowledge obtained from understanding effects of chemicals to ensure that society can use chemicals safely without unacceptable risks to human or animal health [1]. Given the enormous number of already existing chemicals and the ever-increasing number of new chemicals that will be produced and to which we might be exposed, it is unrealistic to assume that we can continue to perform hazard identification, define acceptable use levels, and conduct full risk assessments through animal testing given its limitations (i.e., high numbers, cost, ethics, physiologic differences). Thus, new alternative methodologies (NAMs) hold great promise for making large numbers of

assessments, organizing the results of the assessments, and applying the information from similar chemicals.

Recently, during a roundtable session at the Eurotox Congress (2025) entitled “Advances in Regulatory Science with NAMs and Animal Models” questions were posed and answered by experts, with each question followed by audience participation. Overall, it was a spirited discussion focused on defining NAMs, providing examples of NAMs that have replaced animal models, and identifying the greatest obstacles of adopting NAMs for regulatory use and challenges associated with extrapolating NAMs to human (and animal) health. The goal of this publication is to outline salient points from that discussion with a focus on identifying challenges and opportunities of NAMs development, use, and application for regulatory toxicology. The discussion did not restrict itself to a particular regulatory framework and touched upon diverse EU and national regulations such as the Federal Food, Drug, and Cosmetic Act (FDCA) and the Toxic Substances Control Act (TSCA) in the US, and Registration, Evaluation, Authorisation, and Restriction of Chemicals (REACH) and the regulations for food and cosmetics in the EU.

Beyond the NAM(e): evolution and progress

An initial theme from the session was to focus less on obtaining consensus on the various definitions, including “NAMs” itself or even defining the words in the acronym. Instead, the main points of NAMs are to adopt the principles of the 3Rs, obtain as much relevant information as is possible on chemicals to inform hazard identification and risk assessment, and use this information to protect animal and human health. NAMs include assays performed *in vitro* with cell monolayers or 3-dimensional cellular models (organoids or MPS), *in silico* (QSAR, predictive modeling, AI) read across, or non-mammalian organisms (zebrafish larvae, *C. elegans*). Further discussion focused on physiology in NAMs selection; ensuring the correct assay is being used to address the scientific question and considering what the NAM can deliver. Thus, an important point for successful application of NAMs is a well-constructed and specific context of use; these might be most successful if the NAM addresses a specific question, ideally with a well-understood underlying mechanism or for which key features of toxicity can be reconstructed *in vitro*. Moreover, it is recognized that not all entities, regulatory or otherwise, use the same criteria for “validation”; the process required to ensure the assay indeed answers the intended question with rigor and reproducibility. However, for the purposes of this manuscript, we have used the term “validation” throughout.

The goal for many NAMs is to simulate a human, at least in terms of chemical effects, safety evaluation, and the

practice of regulatory science. Regardless of the intent, for maximum application of the NAM, one needs the cells, organoids, MPS, and/or AI models to reflect the human condition. A particular NAM might need to simulate a particular organ, a function of that organ, and consider the number and types of cells, their differentiation state, and the cell–cell interactions needed to provide normal function, and demonstrate susceptibility to a disease or toxicant exposure. Alternatively, the NAM may need to identify a specific toxicological event, irrespective of the cellular or organ context, such as genotoxicity.

Examples of NAMs being used to replace animal models

A substantial number of NAMs have already found acceptance in regulatory risk assessment (Table 1). However, the type of organ or tissue damage or adversity that they cover varies considerably. For example, irritation of skin and eye are considered by many regulators to be sufficiently covered by NAMs so that data from *in vivo* studies are no longer required. Similarly, the recent OECD-defined approach for skin sensitization has removed much of the need for animal studies to cover this endpoint. Many regulators also accept that in case of clear evidence for no potential for genotoxicity from *in vitro* tests, additional *in vivo* studies are no longer required. In addition, there are several very promising *in vitro* assays for mutagenicity and chromosomal damage under development and evaluation by the OECD. Other endpoints for which NAMs are beginning to make an important impact are identification of endocrine activity and exposure (absorption, distribution, metabolism, and excretion [ADME]). Detailed OECD or ICH test guidelines may only partially cover all the endpoints relevant to endocrine activity or ADME, but they are complemented by a significant body of well-established NAMs.

The main problem in regulatory risk assessment is that systemic toxicity leading to target organ damage, neurotoxicity, reproductive or developmental toxicity, or the development of cancers, remains poorly covered by NAMs. Part of the challenge is that both NAM developers and risk assessors need to address the fact that a NAM cannot be used in isolation to predict many of the apical adverse outcomes observed after repeated and long-term exposure or where complex systemic interactions are needed for adversity manifest (e.g., toxicity resulting from hemodynamic effects).

Obstacles in extrapolating NAMs to human

The use of NAMs can be broadly split into hazard identification and risk assessment. As noted above, NAMs are already successfully applied to hazard identification approaches such as *in vitro* screening during drug discovery to identify molecules with propensities

Table 1 Examples of NAMs already used in regulatory risk assessment. Includes those under OECD evaluation and relevant sources of OECD guidance documents for the application of NAMs.

DERMAL Skin corrosion & skin irritation • <i>In vitro</i> membrane barrier test method for skin corrosion (OECD TG 435) • <i>In vitro</i> skin corrosion: Reconstructed human epidermis (RHE) test method (e.g. EpiSkin™, Epiderm™, SkinEthic™ (OECD TG 431) • Guidance Document on an Integrated Approach on Testing and Assessment (IATA) for skin corrosion and irritation (OECD GD 203) Skin sensitization • Defined Approaches for skin sensitisation (OECD TG 497) • <i>In chemico</i> skin sensitisation and <i>in vitro</i> activation of keratinocytes (OECD TG 442C-E) Skin absorption/penetration • Skin absorption: <i>In vitro</i> method (OECD TG 428) • Dermal absorption (OECD GD 28) OCCULAR CORROSION AND IRRITATION • Reconstructed human cornea-like epithelium (RhCE) test method (OECD TG 492) MUTAGENICITY/GENOTOXICITY Base level testing • Bacterial reverse mutation test (OECD TG 471) • <i>In vitro</i> micronucleus test (OECD TG 487) Other <i>in vitro</i> genotoxicity test methods • <i>In vitro</i> mammalian cell gene mutation tests using the <i>HPRT</i> and <i>XPRT</i> genes (OECD TG 476) • <i>In vitro</i> mammalian cell gene mutation tests using the thymidine kinase gene (OECD TG 490) • <i>In vitro</i> mammalian chromosome aberration test (OECD TG 473) Supportive tests in overall weight of evidence (WoE) approach (for mechanistic understanding) • Pig-a test <i>in vitro</i> (mutation of glycosylphosphatidylinositol [GPI] anchor proteins on the cell surface) • Toxicogenomics (genes involved in DNA instability) • Recombinant cell models (GreenScreen HC, BlueScreen HC, ToxTracker) • γH2AX • Epigenetic responses (e.g., DNA methylation, miRNAs and histone modifications) CARCINOGENICITY No validated alternative <i>in vitro</i> method is available as OECD TGs, but each assay noted adds to a WoE supporting an assessment of carcinogenicity • <i>In vitro</i> Bhas 42 cell transformation assay (Bhas 42 CTA; OECD GD 421; no regulatory acceptance) • Syrian hamster embryo cells transformation assay (SHE CTA; OECD GD 214; no regulatory acceptance) • Toxicogenomics (genes involved in cancer) ENDOCRINE DISRUPTIVE PROPERTIES Several <i>in vitro</i> assays are available (OECD GD 150 [2018]) • Estrogen (OECD TG 493; US EPA TG OPPTS 890.1250) or androgen receptor binding affinity (US EPA TG OPPTS 890.1150, 2009)	Table 1 (continued) • Estrogen receptor transactivation (OECD TG 455; US EPA TG OPPTS 890.1300), human cell-based reporter gene assay (ISO 19040–3:2018), yeast estrogen screen (ISO 19040-1, 19040–2:2018) • Androgen receptor transcriptional activation assay for detection of androgenic agonist and antagonist activity of chemicals (OECD TG 458) • Steroidogenesis <i>in vitro</i> (OECD TG 456 [2022]); US EPA TG OPPTS 890.1550, 2009) • <i>In vitro</i> aromatase assay (US EPA TG OPPTS 890.1200) • Thyroid disruption assays (e.g., thyroperoxidase inhibition, transthyretin binding); a project on validation of selected <i>in vitro</i> methods within EU-NETVAL activity is on-going • Retinoid receptor transactivation assays NEUROTOXICITY AND IMMUNOTOXICITY Limited number of assays available • IL-2 Luc assay test method to evaluate the potential immunotoxic effects of chemicals on T lymphoblastic cell line (OECD TG 444A) • Initial recommendations on evaluation of data from the developmental neurotoxicity (DNT) <i>in vitro</i> testing battery (OECD 2023) REPRODUCTIVE & DEVELOPMENTAL TOXICITY No <i>in vitro</i> assays validated or accepted for regulatory use available for chemicals, guidance for using a NAM to replace a dosage-range embryo/fetal toxicity study provided in ICH S5 R3 SYSTEMIC TOXICITY Acute toxicity • 3T3 NR (neutral red) uptake test (ICCVAM & EURL-ECVAM) • <i>In vitro</i> 3T3 NRU phototoxicity test (OECD TG 432) Sub-acute & sub-chronic toxicity No validated or generally accepted alternative methods available ADME Absorption • Guidance document on the characterisation, validation and reporting of physiologically based kinetic (PBPK) models for regulatory purposes (OECD GD 331) • Harmonised guideline on drug interaction studies (ICH M12) Metabolism & Clearance • Several <i>in vitro</i> assays based on human liver microsomes and hepatocytes • No OECD GD or TG for metabolism or clearance Modelling (PBPK) • Guidance document on the characterisation, validation and reporting of PBPK models for regulatory purposes (OECD GD 331)
---	---

(continued on next page)

to cause organ damage (e.g., drug-induced liver injury, DILI) or adverse effects (e.g., cardiac ion channels) (for these and other examples see Ref. [2]). This supports internal decision making but at present often only supplements regulatory submission. However, risk assessment requires quantitative extrapolation from *in vitro* measures to a human *in vivo* setting. NAM-enabled decision making for early clinical trials is a particular challenge as there is a high need to ensure human safety requiring a comprehensive assessment of systemic toxicity or risks to be able to set safe starting dose and dose escalation levels to avoid side effects. In

particular, capturing human developmental events unfolding over human-scale timing and/or not reaching sufficient maturity is also a limitation of NAMs. Indeed, some biological processes require the recapitulation of the physiological conditions of whole complex biological organism which at present is not possible and will require a series of batteries of NAMs.

The use of human cells is deemed to be an advantage because there is no need to extrapolate across species, but many factors must be considered to mimic all susceptible populations. Different cell types (including cell sex) might be required to interact in a three-dimensional relationship in order to provide prediction of the intact human [3]. The immediate environment of the cells in culture, including cell types (single or multiple origins), the need for specialized media to keep them in a differentiated state (e.g., presence of growth factors, cytokines, and other hormones), and the three-dimensional configuration can all affect outcome measures. The inclusion of multiple cell types and cell layers or scaffolding has proved an important variable in cellular function. One of the primary questions regarding the use of human cells *in vitro* to assess risk is the number of replicates and the number of donors required to provide statistically reliable data. Lin et al., (2023) combined high-throughput human induced pluripotent stem cell (iPSC)-cardiomyocyte (CM) testing with *in silico* data analysis and demonstrated that this *in vitro-in silico* approach could predict proarrhythmic potential with accuracy and precision comparable to other lower-throughput testing methods [4]. The accuracy of the logistic regression model could be improved by using concentration-response data from a population of donors, as compared to using data from single or multiple replicates of a single donor [4].

AI can provide a framework to enhance *in vitro* to *in vivo* extrapolation (IVIVE) with the goal of improving mechanism-based toxicity evaluations [5]. The linking together of several MPS to simulate an intact human-on-a-chip is a possibility with well-constructed and well-tuned fluid dynamic systems. Microfluidic control of multiple organ systems is possible, and models envisioned by bioengineers have been developed. The challenges include the requirement for each organ type to have its own specialized media yet be connected so that the human circulatory system can be replicated. More complex *in vitro* models (e.g., organoids and MPS), face particular challenges with reproducibility due to their complexity [6]. These systems may require precise control of multiple cell types and their ratios, and complex assembly and priming procedures.

Mohapatra et al., (2025) have recently described the core principles for reporting *in vitro* methods and noted that every cell model or test system has unique characteristics, and it is the responsibility of the investigator(s) to

adequately adapt and report these principles in their published studies [7]. The authors outlined their proposed acceptance criteria for NAMs in toxicology, including criteria to evaluate the relevance and reliability of the new method, the importance of aligning these methods with human health-relevant research, and the need for a systematic approach to assess new methodologies. For example, for cell-based *in vitro* models, a systematic approach can refer to quality controls that allow for cell line verification (short tandem repeat analysis, confirming lack of mycoplasma contamination, and avoiding use of xenogenic culture components). The systematic approach can further include documenting culture conditions and maintenance procedures such as passage number, density, and medium change frequency to demonstrate the suitability of the system for use and aid in the comparison of results both within a laboratory and across laboratories [7].

Human embryonic stem cells and iPSCs promise to overcome limitations of traditional cell lines but bring their own challenges that demand specialized quality control measures [8] as differentiation efficiency may differ between batches and extended culture periods increase contamination risk and may cause genetic instability. Issues that need to be addressed include implementation of quality control measures, adoption of standardized reporting guidelines, integration of quality assessment in funding and publication decisions, and development of training programs for good cell culture practices. Some of these limitations can be addressed with the integration of quality standards including automated culture systems to reduce human variability, standardized protocols and reagent preparation, real-time monitoring of culture conditions, contamination detection, and AI-assisted quality control.

Scientific and technical obstacles range from difficulty capturing the complexities of human physiology and toxicology requiring an interplay and exchange between of multiple organs, ability to incorporate all key features into a model to recapitulate a specific bodily/toxicity mechanism, lack of appropriate metabolism and transporter activity to capture ADME features appropriately, to insufficient lifespan of *in vitro* models to recapitulate chronic use in humans. In some cases, it would be relevant to include various microbiomes, for example in a chip that investigates the effects of a test material on the intestine or skin. Perhaps even including parasites known to be common in a human population into an organoid or organ on a chip could also be reasonable and render the model more relevant for extrapolation to humans.

Beyond human extrapolation: other obstacles

More broadly, challenges for the use of NAMs in regulatory decision making include limited regulatory

familiarity with NAMs, lack of confidence in accuracy (acceptable sensitivity and specificity) and predictive power (how well an assay forecasts the actual outcome, whether it be positive or negative) of NAMs, uncertainty about validation requirements (with a range of terminologies used across jurisdictions being a challenge in itself), limited or lack of funding for validation studies, and lack of global harmonized standards. Ultimately collaboration and transparency from both NAM developers as well as regulators will be key to progress and success. There is uncertainty about expectations and standards required for different applications, such as screening versus dose setting in clinical trials, ranging from adequate criteria for fitness for specific context of use, quality control, reliability, reproducibility and predictive power compared with traditional models. There are also ethical considerations, ensuring that the donor(s) have provided permission for use and purpose for their cells. Whilst frameworks and recommendations exist for validation (e.g., OECD, ECVAM, or ICCVAM) [9,10], such studies are lengthy and costly, and lack of funding is a barrier. The requirement for good laboratory practice (GLP) standards is another consideration; whilst this may not be needed for all NAMs, minimum criteria and transparency demonstrating fitness of the model for a given context of use, adequate and justified experimental conditions and controls, performance standards, reference datasets (including choice and number of reference test compounds) and clearly defined applicability domains are critical to underpin the confidence in NAM-derived data. Although the NAM should address a specific question with a well-understood mechanism or key events that can be reconstructed *in vitro*, not all mechanisms and events are characterized enough to develop specific assays capable of detecting them.

Quantitative outcome measures are essential to the overall success of any NAM, and they should reflect the human condition and be relevant to improving the practice of evidence-based medicine and regulation. Most of these quantitative measures should reproduce the normal functioning of the tissue *in situ*, be consistent with a mechanistic understanding and use standard procedures that are widely accepted. Critically, the model system(s) need to reflect the disease state or developmental stage in order to provide a valid prediction over a clinically-relevant exposure range. One recent paper identified that the use of various models can more accurately predict cell concentrations *in vitro* to improve accuracy of IVIVE; the models consider chemical loss to plastics and media [11]. A series of mutually agreed upon positive and negative test agents will be useful to ascertain functional status and reproducibility as well as confirming the applicability domain of the model system. Multi-center trials will be necessary to confirm reproducibility and the

ability to perform the assays in many different environments.

Current efforts on NAMs applications for regulatory use

To relate NAMs applications for regulatory use to our current procedures for conducting safety studies with animals, we need to apply similar principles used to ensure reproducibility and repeatability that are used in animal studies. These procedures were developed over many years because we came to understand many of the issues with using animal models. Procedures that continually ensure consistency of assay results over time need to be made routine for *in vitro* tests as they are for *in vivo* tests. This would include procedures to ensure that methods can be consistently used over time and continue to produce consistent results, while at the same time allowing the NAMs to benefit from technological advances. Ouedraogo et al., in their paper on “a call to action” proposed a unified framework to accept NAMs. The four points of the framework include: 1) establishing the context of use or what a NAM does and does not do; 2) provide standards for any NAM; 3) define or characterize the biological variability and; 4) demonstrate repeatability and reproducibility [12]. Within each part of this framework, we might be able to apply aspects of a systematic approach, including that outlined above [7]. For instance, within “standards for any NAM” involving cell lines, cell line verification assays could be reported when the cells are used regardless of whether the NAM is being used for regulatory decision making or mechanistic determinations. We already consider these four points for our animal models used in a regulatory setting. For example, bio-analytical assays used to determine exposure in a tissue or blood are required to demonstrate reproducibility at least once for each time an assay is used at a specific laboratory. For any NAM to ensure that it is continually acceptable for use involves not only demonstrating that an analytical or bioanalytical assay is reliable and reproducible but also that any organoid, tissue, or cell used had not degraded or mutated over time making the results of any one test not comparable to other test results using the same NAM.

Several regulators covering key regions have published roadmaps for greater use and implementation of NAMs; this reflects the commitment of the community to reduce and phase out animal use for pharmaceutical and chemical testing and other industries where possible and meaningful. Mechanisms supporting such efforts are safe harbor and data sharing proposals or efforts such as the Innovative Science and Technology Approaches for New Drugs (ISTAND) program providing feedback and advice to NAM developers [13]. Similarly, in Europe, the role of the EURL ECVAM Network for Preliminary Assessment

of Regulatory Relevance (PARERE) is to provide input on potential regulatory relevance and suitability of NAMs and testing strategies based on NAMs [14].

The ever-expanding range of pharmaceutical modalities provides both challenges and opportunities for NAMs. There is a long history of use of *in vitro* assays to study and predict toxicities for small molecule drugs and chemicals which may not be appropriate for biologic or cell therapy products requiring the development of new approaches. Conversely, due to small molecules typically having higher probability for promiscuous effects, as they can potentially bind to multiple substrates (e.g., proteins, DNA, RNA), making it more challenging to develop and apply *in vitro* assays to cover all potential risks, whereas biological therapeutics have high specificity for their human targets and adverse effects arise from on-target exaggerated pharmacology. They clearly also highlight scenarios where animal studies are not meaningful based on target expression; pharmacological activity or relevant biology that is not represented in the animal model. This opens the door for innovative approaches. Beilmann et al., recently collated case examples of biologic therapies in which *in vitro* NAMs supported the safety assessment and entry into clinical studies [2]. Whilst these are highly specific therapies and hence the approaches chosen may not be easily transferable to or appropriate for other applications, they still demonstrate clearly that a meaningful safety assessment can be conducted using *in vitro* NAMs only and highlight the potential for NAMs in regulatory decision making allowing much needed therapies to gain regulatory acceptance and entry into clinical studies.

Similarly, authors from FDA/CDER/OND recently presented examples of case studies for NAM data that had been submitted to the agency. Considerations, limitations, and how such data were ultimately used were discussed. This is a highly valuable resource for researchers to understand the requirements for NAM-derived data to be accepted for regulatory decision making and should help to better plan and conduct studies intended to be included in the submission package. However, it also highlighted that NAMs are already used successfully to inform regulators' decisions [15]. Example validation criteria for NAMs for NIH and FDA have been suggested [16].

With the realization that there is utility in AI to enhance safety assessment, there is also a responsibility to maximize the performance of AI approaches. To enhance the effective and efficient use of AI and machine learning (ML) technologies, Tong and colleagues have proposed a framework named TRIAL (transparency, reproducibility, interpretability, applicability

and liability). Distinct from other similar frameworks from academia or other government agencies, the TRIAL framework is intended to serve as a reference point for AI applications that could be adopted across the global regulatory community [17]. More recently, the AI framework for application to regulatory science has been augmented to include five principles: trustworthiness, reproducibility, explainability, applicability and transparency (TREAT) [6,18].

Conclusions

The adoption of NAMs in regulatory toxicology offers a transformative opportunity to reduce animal testing and improve human relevance, but implementation remains complex. While NAMs have gained traction for endpoints such as skin irritation, sensitization, and genotoxicity, significant gaps persist for systemic and chronic toxicity as well as neurotoxicity, developmental, and reproductive toxicity. Scientific challenges include replicating the complexity of human physiology, ensuring mechanistic relevance, and achieving robust quantitative extrapolation from *in vitro* to *in vivo* conditions. Technical obstacles such as reproducibility, standardization of protocols, and quality control for advanced models like organoids and MPS further complicate progress.

Regulatory challenges are equally critical: limited familiarity with NAMs, lack of global harmonization, uncertainty around qualification and acceptance criteria, and inconsistent terminology across jurisdictions hinder confidence and uptake. Resource constraints, including funding for validation studies and the need for multi-center reproducibility trials, add to the burden. Ethical considerations related to donor cell use and practical issues such as variability in cell sources must also be addressed. Emerging technologies such as AI-driven modeling and integrated human-on-chip systems hold promise for overcoming current limitations, but they require rigorous validation and governance. Ultimately, NAMs should be integrated into a WoE approach that aligns with the principles of the 3Rs and advances regulatory science toward greater efficiency, ethical responsibility, and human (and animal) health protection. Moving forward, success will depend on collaborative efforts among regulators, industry, and academia to establish clear contexts of use, harmonized frameworks for acceptance, and transparent performance standards. Indeed, a major hurdle with which we are struggling is the fact that NAMs, whether a single assay or battery of assays, are not a 1:1 replacement of repeated-dose animal observational studies (e.g., a sub-chronic toxicity study) but rather need to be incorporated into more mechanistic, hypothesis-driven frameworks that address key features of toxicity and consider context of use.

Disclaimers

The views presented are those of the author and do not necessarily reflect those of the U.S. Food and Drug Administration or other government agency.

Generative AI

Generative AI (Microsoft Co-Pilot) was used to create the initial versions of the abstract and conclusion sections to which all authors made comments and edits prior to submission.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Malvezzi S, Buche C, Lorez C, Raybould A: **From basic research in toxicology to a career in regulatory toxicology: an industry perspective.** *Chem Res Toxicol* 2019, **32**: 1901–1903.
2. Beilmann M, Adkins K, Boonen HCM, Hewitt P, Hu W, Mader R, Moore S, Rana P, Steger-Hartmann T, Villenave R, et al.: **Application of new approach methodologies for nonclinical safety assessment of drug candidates.** *Nat Rev Drug Discov* 2025, **24**:705–725.
- This review summarizes progress for NAMs in drug development and encourages use of NAMs for human-relevant safety assessment.
3. Slikker Jr W: **Of human-on-a-chip and humans: considerations for creating and using microphysiological systems.** *Exp Biol Med (Maywood)* 2014, **239**:1078–1079.
4. Lin HC, Rusyn I, Chiu WA: **Assessing proarrhythmic potential of environmental chemicals using a high throughput in vitro-in silico model with human induced pluripotent stem cell-derived cardiomyocytes.** *ALTEX* 2024, **41**:37–49.
5. Chandra M, Li T, Tong W: **AIVIVE: a novel AI framework for enhanced in vitro to in vivo extrapolation (IVIVE) of toxicogenomics data.** *Toxicol Sci* 2025, **207**:361–371.
- This manuscript shows how AI can enhance interpretation of already-available toxicogenomics data.
6. Hartung T, Whelan M, Tong W, Califf RM: **Is regulatory science ready for artificial intelligence?** *npj Digit Med* 2025, **8**:200.
7. Mohapatra R, Leist M, von Aulock S, Hartung T: **Guidance for good in vitro reporting standards (GIVReSt) - a draft for stakeholder discussion and background documentation.** *ALTEX* 2025, **42**:376–396.
8. Pamies D, Ekert J, Zurich MG, Frey O, Werner S, Piergiovanni M, Freedman BS, Keong Teo AK, Erfurth H, Reyes DR, et al.: **Recommendations on fit-for-purpose criteria to establish quality management for microphysiological systems and for monitoring their reproducibility.** *Stem Cell Rep* 2024, **19**: 604–617.
9. ECVAM E. 2025. https://joint-research-centre.ec.europa.eu/projects-and-activities/reference-and-measurement/european-union-reference-laboratories/eu-reference-laboratory-alternatives-animal-testing-eurl-ecvam_en.
10. ICCVAM: **Validation, qualification, and regulatory acceptance of new approach methodologies.** 2024. <https://ntp.niehs.nih.gov/whatwestudy/niceatm/resources-for-test-method-developers/submissions>.
11. Scherer MN, Feshuk M, Armitage JM, Arnot JA, Crizer D, DeVito M, Ferguson SS, Freeman K, Honda GS, Harrill JA, et al.: **Characterizing accuracy of model predictions for chemical concentration in high throughput screening assays.** *Environ Sci Technol* 2026, **60**:4869–4877.
12. Ouedraogo G, Alepee N, Tan B, Roper CS: **A call to action: advancing new approach methodologies (NAMs) in regulatory toxicology through a unified framework for validation and acceptance.** *Regul Toxicol Pharmacol* 2025, **162**, 105904.
- This paper provides a historical overview of NAMs development and provides specific examples use of NAMs in various industries.
13. FDA U. <https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/innovative-science-and-technology-approaches-new-drugs-stand-program>; 2025.
14. ECVAM E. https://joint-research-centre.ec.europa.eu/projects-and-activities/reference-and-measurement/european-union-reference-laboratories/eu-reference-laboratory-alternatives-animal-testing-eurl-ecvam/alternative-methods-toxicity-testing/advisory-and-consultation-bodies/parere-eurl-ecvam-network_en; 2026.
15. Yao J, Peretz J, Bebenek I, Avila A, Alapatt T, Lee B, Patel D, Brown P, Davis-Bruno K: **FDA/CDER/OND experience with new approach methodologies (NAMs).** *Int J Toxicol* 2025, **109**15818251384270.
16. Tieghi RS, Hoeg TB, Makary MA, Kleinstreuer N: **The promise of animal testing alternatives at the U.S. food and drug administration and national institutes of health.** *New Engl J Med AI* 2025, **3**:1.
17. Anklam E, Bahl MI, Ball R, Beger RD, Cohen J, Fitzpatrick S, Girard P, Halamoda-Kenzaoui B, Hinton D, Hirose A, et al.: **Emerging technologies and their impact on regulatory science.** *Exp Biol Med (Maywood)* 2022, **247**:1–75.
18. Tong W, Renaudin M, Taskforce GIL: **Context is everything in regulatory application of large language models (LLMs).** *Drug Discov Today* 2024, **29**, 103916.