Exemplified Framework for Replacing Animal Testing in Challenging Toxicology Endpoints Using New Approach Methodologies (NAMs)

A technical report from the Alliance for Cruelty Free Science

Date: January 2026

Exemplified Framework for Replacing Animal Testing in Challenging Toxicology Endpoints Using New Approach Methodologies (NAMs)[1]

A Technical Report from the Alliance for Cruelty Free Science

Date: January 2026

This document is freely available for use to support the drive of global progress toward animal-free toxicology as soon as possible.

[1] In this context New Approach Methodologies =100% animal free replacement methods not complementary to continued use of animals in anyway.

Contents:

Executive Summary

As of late 2025, global regulatory momentum to phase out animal testing in toxicology has reached a pivotal point. The U.S. FDA's *Roadmap to Reducing Animal Testing in Preclinical Safety Studies*[2] (published April 2025) outlines a strategic, stepwise approach to make animal studies the "exception rather than the norm," beginning with monoclonal antibodies and leveraging NAMs such as organ-on-chips, computational modeling, and advanced in vitro assays. In parallel, ICCVAM[3]/ NICEATM[4] is advancing the CAMERA database[5] (beta version mid-2025) as a centralized resource for validated NAMs, while the OECD[6] continues to develop IATA[7] case studies for complex endpoints.

Despite this progress, several "hard-to-replace" endpoints—

- repeated-dose systemic toxicity,
- reproductive/developmental toxicity (DART),
- carcinogenicity, inhalation toxicity,
- neurotoxicity (including developmental neurotoxicity),
- immunotoxicity,
- ecotoxicology

remain reliant on animal studies due to challenges in the problem of getting cutting edge science solutions, validated.

This open-source exemplified white paper, prepared by the Alliance for Cruelty Free Science, suggests how to addresses these gaps. It suggests a tiered Integrated Approaches to Testing and Assessment (IATA) built entirely on human-relevant NAMs. Drawing from Next Generation Risk Assessment (NGRA)[8] principles—exposure-led, hypothesis-driven, and integrating in silico, in vitro (e.g., multi-organ chips, organoids), and omics data—these IATA provide structured decision trees for hazard classification, point-of-departure (PoD) derivation, and risk assessment without defaulting to animal data.

---

[2] https://www.fda.gov/files/newsroom/published/roadmap_to_reducing_animal_testing_in_preclinical_safety_studies.pdf

[3] https://ntp.niehs.nih.gov/whatwestudy/niceatm/iccvam

[4] https://ntp.niehs.nih.gov/whatwestudy/niceatm

[5] https://camera.niehs.nih.gov/

[6] https://www.oecd.org/en.html

[7] https://www.oecd.org/en/topics/sub-issues/assessment-of-chemicals/integrated-approaches-to-testing-and-assessment.html

[8] Next Generation Risk Assessment (NGRA) is a modern, human-relevant approach to chemical safety that uses New Approach Methodologies (NAMs)—like computational models, in vitro tests, and in chemico assays—to assess harm, moving away from animal testing by focusing on exposure-led, hypothesis-driven science.

Here's why we see it as valuable and potentially novel:

1. **Filling Implementation Gaps:**

While major players like the FDA's April 2025 Roadmap, OECD's IATA case studies, and HESI's reports[9] provide high-level strategies and validation pathways, they often stop short of fully exemplified, endpoint-specific decision trees with visual flowcharts and tiered structures tailored for immediate regulatory use.

Our white paper explicitly suggests how to bridge this by:

- proposing practical IATA built on NGRA principles,
- integrating tools like organoid batteries,
- AI PK/PD modeling[10],
- IVIVE[11]

We also have suggested areas to add to an implementation roadmap (short/mid/long-term pilots, harmonization, and training). This hands-on approach could help regulators and industry move from theory to proactive submissions, aligning with but extending beyond broader roadmaps.

Implementation Roadmap

- **Short-term**: Pilot for specific classes.
- **Mid-term**: Regulatory harmonization.
- **Long-term**: Full replacement, training programs.

2. Open-Source and Collaborative Nature:

Released under CC BY 4.0, this technical report invites discussion, adaptation and use among the wider community interested in and concerned with the issues of difficult to replace toxicology endpoints for the use of animals in science.

This approach isn't always possible or viable in the case with agency or industry documents (e.g. some FDA guidances are more prescriptive without encouraging direct builds). This white paper calls on stakeholders to "use and build upon" it, to help them position catalysts for global progress.

This is a discussion and reference report for those looking to develop a regulatory framework that can be shared.

---

[9] HESI reports refer to the detailed score reports and performance analytics generated from exams produced by Health Education Systems Incorporated (HESI), a company owned by Elsevier that specializes in assessments for nursing and health professions education.

[10] AI PK/PD modeling refers to the application of artificial intelligence (AI) and machine learning (ML) techniques to pharmacokinetics (PK) and pharmacodynamics (PD) modeling in drug development and precision medicine.

[11] IVIVE stands for In Vitro to In Vivo Extrapolation, a key methodology in pharmacology, toxicology, and drug development that bridges data from laboratory (in vitro) experiments to predictions of behavior in living organisms (in vivo).

3. Focus on Human-Relevant, Ethical Advances:

The emphasis on weight-of-evidence integration[12] (AOP-informed, with uncertainty characterization) and comparisons to traditional testing highlighting NAMs' superior human relevance and speed) adds a clear, evidence-based narrative.

It also ties into emerging trends like Certara's weight-of-evidence philosophy[13] but this white paper applies specifically to those hard-to-replace endpoints identified.

This paper could differentiate it from more general NAM overviews, offering an opportunity for industry and regulators to develop tools to confidently remove animal studies without sacrificing safety, through active collaboration.

Key features:

- **Tiered structure**: From exposure-based waiving and in silico predictions to advanced bioactivity screening and IVIVE (in vitro to in vivo extrapolation) via PBK modeling[14].
- **Weight-of-evidence integration**: AOP-informed[15], with uncertainty characterization.
- **Visual decision flowcharts**: For regulatory usability
- **Implementation roadmap**: Short-term pilots, training, and harmonization with FDA, ECHA[16], and OECD.

This suggested framework helps to bridge regulatory expertise gaps, builds confidence in NAMs, and accelerates the ethical transition to more predictive, human-relevant science—ultimately sparing millions of animals while enhancing safety decisions.

The Alliance for Cruelty Free Science calls on regulators, industry, and scientists and the industry to collaboratively build upon the suggestions and ideas in this white paper to end unnecessary animal suffering in toxicology.

---

[12] Weight-of-Evidence (WoE) Integration refers to a structured, systematic process used primarily in toxicology, risk assessment, and regulatory science to evaluate and combine multiple lines of evidence (LoEs) from diverse sources to reach a defensible conclusion about a scientific question, such as hazard identification, causality, or risk characterization.

[13] https://www.certara.com/on-demand-webinar/the-pursuit-of-certainty-integrated-evidence-and-the-evolving-drug-development-and-reimbursement-landscape/

[14] PBK (Physiologically Based Kinetic/Pharmacokinetic) modeling uses mathematical equations to simulate how a chemical or drug is **A**bsorbed, **D**istributed, **M**etabolized, and **E**xcreted (ADME) in the body, representing it as interconnected organs. These models predict internal tissue concentrations from external exposure, helping assess chemical safety, understand dose-response, enable in vitro-to-in vivo (IVIVE) extrapolation, reduce animal testing, and set safe exposure levels for humans and animals.

[15] "AOP-informed" generally refers to using an **Adverse Outcome Pathway (AOP)** framework, a scientific tool that maps the chain of biological events from a chemical's initial interaction with a molecule to a harmful outcome, to guide testing, organize data, and predict chemical risks for human health and the environment, often integrating new testing methods. It's about understanding the 'how' and 'why' of toxicity, rather than just 'what' happens, to make better regulatory decisions, especially when traditional tests are limited.

[16] ECHA, the European Chemicals Agency, is the EU body that manages chemical safety, implementing regulations like REACH to protect health and the environment by overseeing chemical registration, evaluation, authorization, and restriction, making it Europe's central source for chemical information and ensuring safe use by industry and consumers. It handles rules for chemicals, biocides, and persistent organic pollutants, providing databases and guidance for compliance.

Explanation of New Approach Methodologies (NAMs) in Toxicology: Regulatory Acceptance Status

New Approach Methodologies (NAMs) are modern, non-animal-based techniques used in toxicology to assess chemical safety and potential health risks. These include computational tools, cell-based assays, organ models, and mechanistic frameworks. They aim to replace, reduce, or refine traditional animal testing while providing more human-relevant data. Regulatory agencies like the FDA (U.S.), EMA (Europe), and OECD (international guidelines) are increasingly supporting NAMs through roadmaps, qualification programs, and guidance, driven by laws like the FDA Modernization Act 2.0 and EU strategies for sustainability.

The table shared summarizes the **most advanced and widely accepted NAMs** as of 2025, based on their status with major regulators. It ranks them by maturity, industry use, and acceptance level. This is based on available published information at the time of writing this paper.

Here's a clear breakdown:

Highlights of the Most Advanced NAMs

1. **In Silico Methods (e.g. QSAR, Machine Learning, Read-Across)** Most mature and routinely accepted for initial screening (e.g., genotoxicity, skin sensitization). FDA, EMA, and OECD fully endorse them under validated guidelines. Widely used in industry for quick predictions without lab work.
2. **Adverse Outcome Pathways (AOP) and Integrated Approaches to Testing and Assessment (IATA)** Core frameworks for combining data (mechanistic + computational + in vitro). Over 150 AOPs endorsed by OECD; FDA and EMA integrate them into decision-making. Very high adoption as they provide scientific confidence without always needing new animal tests.
3. **Liver-on-Chip and 3D Hepatic Spheroids** Advanced organ models for predicting drug-induced liver injury (DILI). Liver-chips are in FDA qualification pilots (e.g., Emulate's model accepted in ISTAND program). High industry use, especially for human-relevant toxicity testing.
4. **Other Cell-Based Assays (e.g., T-Cell Activation, Tissue Cross-Reactivity)** Encouraged or expected for specific uses like biotherapeutics (e.g., monoclonal antibodies). Mature for targeted screening.

Overall Trends

- **Mature NAMs** (routine green checks across agencies) are already replacing animal tests in many cases, especially for screening.
- **Advanced/Emerging** ones (like organ-chips) are gaining traction through qualification pathways and case studies.
- **Industry adoption** is highest for cost-effective, high-throughput methods (in silico, AOP/IATA).
- **Regulators** are collaborating (e.g., FDA-EMA workshops) to harmonize acceptance, with ongoing efforts to build databases and standards.
- FDA's roadmap aims to make animal testing "the exception" for biologics, expanding to chemicals.
- EMA and OECD support qualification; databases like CAMERA emerging.

- Full replacement delayed by gaps in chronic/systemic toxicity and validation hurdles.

Summary Table Breakdown

Key Columns Explained

- **NAM Type**: The specific methodology.
- **FDA Status (2025)**: Level of support or acceptance by the U.S. Food and Drug Administration.
- **EMA Status (2025)**: European Medicines Agency perspective.
- **OECD Status (2025)**: Organisation for Economic Co-operation and Development (harmonized international guidelines).
- **Industry Adoption**: How widely used in pharma/chemical companies.
- **Maturity Level**: Overall stage (Mature = routine use; Advanced = growing evidence; Emerging = promising but limited).
- **Legend**: Green check = Accepted/Routine; Yellow = Encouraged/Expected; Red dot = Limited.

NAM Type	FDA Status (2025)	EMA Status (2025)	OECD Status (2025)	Industry Adoption	Maturity Level
In Silico (QSAR/ ML, Read-Across)	Accepted (Tier 1 screening)	Accepted	Accepted	Very High (>90% pharma)	Mature
Computational (PBK/PBPK)	Encouraged	Encouraged	Endorsed	High (growing)	Advanced
Liver-on-Chip (DILI)	Encouraged active; pursuing qualification	Qualification pathway active	Case examples published	High (65.7% of organ-chip pubs)	Mature
3D Hepatic Spheroids	Widely accepted (supporting evidence)	Accepted	Accepted	Very High	Mature
Human T-Cell Activation	Encouraged for biotherapeutics	Encouraged	Emerging	High (biotherapeutics)	Advanced
Tissue Cross-Reactivity (TCR)	Expected for mAbs/ bispecific	Expected	Expected	Very High	Mature
Cell-Based Protein Arrays (CBPA)	Expected for off-target screening	Expected	Endorsed	High	Mature
Adverse Outcome Pathways (AOP)	Integrated into IATA; core framework	Core framework	>150+ endorsed AOP	Very High	Mature
IATA (Weight-of-Evidence, QbD)	FDA strategy incorporates IATA	Core acceptance mechanism	Guidance documents published	Very High	Mature

This shift supports better human prediction, ethical goals (reducing animal use), and faster drug/ chemical development.

Comparison: New Approach Methodologies (NAMs) vs. Traditional Animal Testing in Toxicology

New Approach Methodologies (NAMs) encompass non-animal techniques like in vitro (cell-based) assays, organ-on-a-chip models, 3D spheroids, in silico computational modeling (e.g., QSAR, AI), and integrated approaches (e.g., Adverse Outcome Pathways). These are increasingly promoted to replace, reduce, or refine animal use under the 3Rs principle. Traditional animal testing relies on in vivo studies, primarily in rodents (mice/rats), dogs, or non-human primates, observing whole-body effects like organ damage or behavioral changes.

As of late 2025, regulators (FDA, EMA, OECD) are accelerating NAM adoption via roadmaps and pilots, but animal testing remains required for many endpoints due to historical reliance and challenges in full validation.

Key Comparison Table

Aspect	Traditional Animal Testing	New Approach Methodologies (NAMs)
Human Relevance	Often limited due to species differences; ~90% of drugs passing animal tests fail in humans (poor translation for toxicity/efficacy).	Higher; uses human cells/tissues, better predicts human responses (e.g., drug-induced liver injury).
Ethical Considerations	Involves animal suffering; raises welfare concerns.	Animal-free; aligns with 3Rs (replace/reduce/refine).
Speed & Cost	Slow (months/years) and expensive (animal housing/care).	Faster (days/weeks) and cheaper; high-throughput screening possible.
Mechanistic Insight	Observes apical endpoints (e.g., tumors, death) but often lacks deep mechanisms.	Provides detailed molecular/pathway data (e.g., via omics).
Complexity Coverage	Captures whole-organism interactions (e.g., systemic, chronic, multi-organ effects).	Limited for full systemic/repeated-dose toxicity; integration needed for complex endpoints.
Predictivity for Idiosyncratic Toxicity	May miss rare human reactions; mechanisms differ across species.	Challenging for rare events; better for direct toxicity but not yet proven for all idiosyncrasies.
Regulatory Acceptance (2025)	Gold standard; required for many submissions. Yet contiunes to be an unvalidated process without empirical proof.	Growing (e.g., FDA roadmap phases out for some biologics; OECD endorses several); mature for screening (e.g., skin sensitization) but not full replacement for systemic toxicity.
Validation & Reproducibility	Established guidelines but reproducibility issues across labs/species.	Needs more standardization; some overly conservative predictions.

Most Difficult Areas for Regulatory Acceptance of NAMs in Toxicology (as of December 2025)

While New Approach Methodologies (NAMs) have achieved strong regulatory acceptance for **local/single endpoints** (e.g., skin sensitization, eye irritation, genotoxicity screening via in silico/QSAR and in vitro assays), several systemic and complex endpoints remain challenging. These areas still largely rely on traditional animal testing due to scientific, technical, and validation hurdles. Regulators (FDA, EMA, OECD, ECHA) acknowledge progress but emphasize the need for more data, integration (e.g., via IATA/AOP frameworks), and confidence-building through pilots and case studies.

Key Difficult Endpoints and Reasons

Endpoint	Difficulty Level	Main Reasons for Regulatory Challenges	Current Status/Examples
Repeated-Dose/ Chronic Systemic Toxicity	Highest	- Captures long-term, multi-organ, cumulative effects and whole-body interactions. - NAMs (e.g., organ-chips, MPS) good for short-term/acute but struggle with prolonged exposure, metabolism, and systemic distribution. - Lack of fully validated integrated models for toxicokinetics/dynamics over weeks/months.	- Major focus of FDA/EMA roadmaps and EPAA/EURL ECVAM projects. - Databases building mechanistic knowledge, but no full replacement yet. - Often requires hybrids (NAMs + limited animals).
Reproductive & Developmental Toxicity (DART/Dev Tox)	Very High	- Involves dynamic multi-generational processes, critical windows, and complex embryogenesis. - Current NAMs (e.g. zebrafish, organoids) useful for screening but not predictive enough for full replacement. - Need robust comparative data to in vivo embryo-fetal studies.	- HESI DART surveys highlight screening use; regulators seek equivalence data. - EURION/ASPIS clusters advancing, but validation slow.
Respiratory Sensitization	High	- No widely accepted AOP or validated in vivo/in vitro methods. - Difficult to model immune-mediated respiratory responses.	- Contrasts with skin sensitization (fully NAM-replaced in many cases). - Proposed AOPs exist but not regulatory-ready.
Neurotoxicity & Developmental Neurotoxicity (DNT)	High	- Complex brain development, blood-brain barrier, and behavioral endpoints hard to replicate. - In vitro batteries emerging but lack systemic integration.	- ECHA/REACH highlights challenges; ongoing OECD/EURL ECVAM efforts.
Carcinogenicity	High	- Long-term, multi-step process involving chronic exposure and genomic instability. - NAMs (e.g., cell transformation assays) supportive but not standalone.	- Partial use in weight-of-evidence; full animal studies often required.
Full Systemic/ Immunotoxicity	High	- Complex immune responses, idiosyncratic reactions, and multi-organ crosstalk. - NAMs conservative but miss rare events.	- Encouraged for biotherapeutics; gaps in chronic assessment.

Why These Are Difficult Overall

- **Scientific Complexity**: These endpoints involve **whole-organism integration** (e.g., ADME[17], immune modulation, chronic adaptation) that current NAMs (mostly in vitro or in silico) cannot fully mimic without advanced multi-organ systems or AI-enhanced modeling.

- **Validation & Confidence**: Regulators require demonstrated **human relevance, reproducibility, and predictivity** equivalent (or superior) to animal data. Slow pace of qualification (e.g., only ~8 NAMs in FDA's ISTAND pilot by mid-2025).

- **Regulatory Inertia**: Historical reliance on animal "gold standard"; need for harmonization (ICH, OECD) and case-by-case acceptance.

- **Data Gaps**: Limited long-term datasets; ongoing efforts (e.g., CAMERA database, EU-ToxRisk) to build evidence.

- **Progress Drivers**: FDA's 2025 roadmap prioritizes these via pilots; EMA/ECHA focus on IATA; international collaborations accelerating.

In contrast, mature areas like **in silico screening** and **local toxicity** are routinely accepted.

The shift is gradual.

NAMs are increasingly used for **early screening, weight-of-evidence, or waiving animal tests** in specific contexts, with full replacement expected first in targeted domains (e.g., biologics).

By 2026 expect more pilots and guidance, but these complex endpoints will likely remain the bottlenecks for years unless some key frameworks are agreed and shared.

The time is now.

---

[17] **ADME** is an acronym widely used in **pharmacology**, **pharmacokinetics**, and **drug development**. It stands for:

- **Absorption**: How the drug or chemical enters the bloodstream (e.g., via oral intake, injection, or skin).
- **Distribution**: How the drug spreads throughout the body to tissues and organs.
- **Metabolism**: How the body chemically breaks down or transforms the drug (primarily in the liver).
- **Excretion**: How the drug and its metabolites are eliminated from the body (e.g., via urine or feces).

These processes determine a drug's bioavailability, efficacy, duration of action, and potential toxicity. ADME studies are essential in early drug discovery to predict how a compound behaves in the body and to identify candidates likely to succeed in clinical trials.

Sometimes it's extended to **ADMET**, adding **Toxicity**, to include safety assessments.

Key Building Blocks for the Framework

A strong framework should be **tiered**, **context-of-use specific**, and aligned with emerging global thinking (e.g., ICCVAM's modernized validation approach, OECD IATA, FDA's Predictive Toxicology Roadmap). These elements have already been largely agreed if not comprehensively used as a regulatory reference globally.

Core elements to include:

1. **Defined Approaches (DAs) and Integrated Approaches to Testing and Assessment (IATA)** Combine multiple NAMs (e.g., in vitro assays + QSAR/read-across + PBK modeling + omics) into structured decision trees. OECD has templates for this—your framework could propose specific IATA case studies for each hard endpoint.
2. **Validation and Confidence-Building Criteria** Draw from recent proposals (e.g., 2025 papers calling for unified, measurable standards: reproducibility, mechanistic plausibility, applicability domain, uncertainty characterization). Include performance metrics vs. historical animal/human data.
3. **Data Requirements and Sources** Emphasize open datasets (e.g., ECHA's IUCLID FDA toxicity data migration, ToxValDB, COSMOS). Propose minimum information standards for NAM submissions.
4. **Weight-of-Evidence (WoE) Integration** How to combine NAM outputs with limited existing animal data or human evidence.
5. **Acceptance Pathways**
    - Tiered testing: Start with high-throughput screening → targeted in vitro/MPS → in silico refinement.
    - Waivers/reductions where NAMs show equivalence or superiority.
    - Pilot acceptance for specific chemical classes (e.g., pesticides, pharmaceuticals).
6. **AI/ML Integration** Guidelines for transparent, explainable models (e.g., Good Machine Learning Practice principles).
7. **Roadmap for Implementation** Short/mid/long-term goals, training needs (like NURA program), international harmonization.

Existing Foundations to Build On

- **FDA Roadmap (2025)**: "Roadmap to Reducing Animal Testing in Preclinical Safety Studies" – emphasizes NAMs for drugs/biologics.
- **ICCVAM/NICEATM Efforts**: Acute toxicity strategy; upcoming CAMERA database (beta mid-2025) for validated NAMs; workgroups on complex endpoints.
- **OECD Guidance**: Defined Approaches for skin sensitization (already accepted); expanding to others.
- **2025 Publications**: Calls for unified validation frameworks (e.g., in ALTEX, Regulatory Toxicology and Pharmacology).
- **EU/EPAA**: Push for NGRA (Next Generation Risk Assessment).

References and Citations

1. U.S. FDA. (2025). *Roadmap to Reducing Animal Testing in Preclinical Safety Studies*. https://www.fda.gov/media/186092/download
2. ICCVAM/NICEATM. (2025). CAMERA database initiatives.
3. OECD. (2025). IATA case studies and guidance updates.
4. Dent et al. (2018-2025 updates). ICCR principles of NGRA; recent applications in systemic toxicity.
5. Various reviews (2025): NAMs for DART, carcinogenicity, etc. (e.g., Archives of Toxicology, Frontiers in Toxicology).

Hardest-to-Replace Endpoints/Areas

Here's a breakdown of the most commonly cited challenging areas, based on regulatory discussions (FDA, EPA, ECHA/REACH), expert panels, and recent reviews:

1. **Repeated-Dose Systemic Toxicity (including Chronic Toxicity)** The "gold standard" 90-day or 2-year rodent studies capture cumulative effects across organs. NAMs can predict acute hits well, but modeling prolonged low-dose exposure and adaptive responses is limited.

2. **Reproductive and Developmental Toxicity (DART)** Involves complex processes like gametogenesis, implantation, organogenesis, and postnatal effects. No fully validated NAM battery yet replicates the full in utero environment and multi-generational impacts.

3. **Carcinogenicity** 2-year bioassays detect tumor formation from chronic exposure. In vitro/in silico approaches identify genotoxins reliably, but non-genotoxic carcinogens and promoter effects are tough without long-term whole-organism data.

4. **Inhalation Toxicity (Respiratory Tract)** Modeling aerosol deposition, lung clearance, and systemic uptake after inhalation is complex. Some in vitro airway models exist, but they don't fully capture deep lung or repeated-exposure dynamics.

5. **Acute Systemic Toxicity (Oral/Inhalation, including Lethality Endpoints)** While reduced for pharmaceuticals, still required for chemicals/pesticides under REACH/OECD. Predicting LD50/LC50[18] without animals remains incomplete due to variability in metabolism and multi-organ failure modes.

6. **Neurotoxicity and Behavioral Endpoints** Assessing developmental neurotoxicity, seizures, or cognitive/behavioral changes requires integrated nervous system responses that current NAMs (e.g., neural organoids) only partially mimic.

7. **Immunotoxicity (beyond Skin Sensitization)** Systemic immune responses, immunosuppression, or autoimmunity involve dynamic whole-body interactions hard to replicate in vitro.

8. **Ecotoxicology (Environmental Impact)** Effects on wildlife (fish, birds, bees) often require species-specific whole-organism tests; aquatic models have alternatives, but terrestrial/mammalian-like endpoints lag.

[18] LD50 and LC50 are standard measures of acute toxicity in toxicology, used to assess how poisonous a substance is.

Detailed Endpoint-Specific IATA Frameworks

Each section below provides a detailed tiered IATA, drawing on current NAMs and regulatory progress (e.g., OECD IATA guidance, FDA Roadmap, NGRA principles).

1. Outline for a Repeated-Dose Systemic Toxicity IATA Framework Using New Approach Methodologies (NAMs)

This outline proposes an exemplified Integrated Approach to Testing and Assessment (IATA) for repeated-dose systemic toxicity (e.g., targeting 28-day/90-day sub chronic endpoints like those in OECD TG 407/408). It draws from emerging 2025 frameworks, such as tiered NGRA approaches (e.g., ECETOC-inspired classification systems), physiologically-based kinetic (PBK) modeling integrated with in vitro bioactivity, and ongoing ICCVAM/OECD/EPA/FDA efforts to prioritize NAMs for systemic endpoints.

The goal is a non-animal, human-relevant, tiered decision framework that provides points of departure (PoDs) for hazard identification, classification (e.g., STOT-RE[19]), and risk assessment, while building regulatory confidence through weight-of-evidence (WoE).

The purpose is to provide regulatory bodies with a suggested framework for hard to replace areas of animal testing as of December 2025.

1. Introduction and Scope

Purpose: To enable full replacement of in vivo repeated-dose studies (e.g., rodent 90-day) with NAM batteries for industrial chemicals, pesticides, pharmaceuticals, and cosmetics.

Applicability Domain**:** Oral route initially; extension to inhalation/dermal. Focus on systemic effects (liver, kidney, nervous system, hematology, etc.); excludes local effects.

Regulatory Alignment: Compatible with OECD IATA guidance, FDA Predictive Toxicology Roadmap (2025 updates emphasizing NAMs for preclinical safety), ICCVAM acute/chronic strategies, and EU NGRA principles.

Decision Context: Hazard classification (low/medium/high concern per recent proposals), NOAEL/LOAEL [20]estimation, or waiver for higher-tier animal testing.

[19] STOT-RE is a **hazard classification** used in chemical safety regulations (e.g., under the UN's Globally Harmonized System - GHS, adopted by EU CLP, OSHA in the US, and many countries worldwide). It identifies substances that cause **damage to specific organs** (e.g., liver, kidneys, nervous system, lungs) after **repeated or prolonged exposure**, even if no effects are seen from a single dose.

[20] NOAEL/LOAEL estimation is a key part of toxicology testing (also called toxicity studies or safety testing). These tests determine how safe a substance is for humans, animals, or the environment. They are required by regulators like the EPA (USA), EFSA (Europe), or WHO before a product can be approved.

2. Principles and Confidence-Building

Tiered Structure: Progressive information gathering to minimize testing while maximizing predictivity.

Uncertainty Characterization: Applicability domain, reproducibility, mechanistic relevance (per OECD GD 34 updates).

Validation Approach: Performance vs. historical animal/human data; use of open databases (e.g., ToxValDB, COSMOS, upcoming CAMERA beta mid-2025).

WoE Integration: Bayesian or structured expert judgment for combining data sources.[21]

21

Overview of Bayesian Methods and Structured Expert Judgment (SEJ) for Combining Data Sources

In fields like risk assessment, toxicology, environmental science, and decision-making under uncertainty, combining data from multiple sources (e.g., experimental studies, observational data, simulations, or sparse empirical evidence) is essential. Two prominent approaches are **Bayesian methods** and **structured expert judgment (SEJ)**. These can be used separately or integrated, depending on data availability and the need to quantify uncertainty. Bayesian methods excel in probabilistic data integration, while SEJ is particularly useful when empirical data is limited, relying on calibrated expert opinions as a surrogate or complement to data.

Both approaches are applied in regulatory contexts, such as by the EPA, WHO, or in nanomaterial risk screening, to derive parameters like toxicity thresholds

3. Tiered IATA Workflow

Tier 0: Exposure and Prioritization (No Testing Needed for Low-Exposure Scenarios)

- Use Threshold of Toxicological Concern (TTC) or exposure-based waiving.
- Read-across/category approach from analogs (e.g., OECD QSAR Toolbox).

Outcome: Low concern → no further testing; proceed otherwise.

Tier 1: In Silico Predictions and Toxicokinetics (TK)

- Physicochemical properties and ADME predictions (e.g., OPERA, VEGA tools).
- PBK/PBTK modeling for internal dosimetry (e.g., httk R package; reverse dosimetry to estimate human equivalent doses).
- (Q)SAR for repeated-dose endpoints (e.g., CATMoS models for NOAEL prediction).
- Bioavailability assessment (e.g., in vitro Caco-2 permeability).

Outcome: Predict bioaccumulation/chronic potential; refine exposure; flag high-concern structures.

Tier 2: High-Throughput In Vitro Bioactivity Screening

- ***Broad coverage assays:*** ToxCast/Tox21 battery (ER, AR, PPAR, etc.); high-content imaging for cytotoxicity.
- ***Targeted organ-specific assays:***
  - Liver: HepaRG/primary hepatocytes (steatosis, cholestasis via omics).
  - Kidney: RPTEC models (tubular toxicity).
  - Neuro: Neural organoids or iPSC-derived models.
  - Multi-organ chips (MPS) for inter-organ crosstalk (e.g., liver-kidney-lung systems).
- ***Repeated-exposure simulation:*** Dose fractionation over 14-28 days in vitro.
- ***PoD derivation:*** Benchmark concentration (BMC) from concentration-response curves.

Outcome: Bioactivity PoD (e.g., lowest AC50 across assays); classify as inactive/low/medium/high potency.

Tier 3: Advanced Integration and Refinement

IVIVE (In Vitro to In Vivo Extrapolation): Convert in vitro PoDs to external doses via PBK (accounting for repeated dosing, steady-state Css).

- ***AOP-informed WoE:*** Link bioactivity to key events (e.g., via AOP-Wiki).
- ***Omics integration:*** Transcriptomics/metabolomics for mode-of-action (e.g., TempO-Seq for short-term in vivo extrapolation to chronic).
- ***Uncertainty factors:*** Apply based on coverage gaps (e.g., 3-10 for extrapolation).

Outcome: Human-relevant PoD (e.g., equivalent NOAEL in mg/kg/day); classification into low/medium/high concern.

Tier 4: Case-Specific Targeted Testing (If Needed)

Additional NAMs only if critical gaps (e.g., advanced multi-organ chips, emerging omics, or enhanced read-across from larger analog sets); otherwise, rely on existing human epidemiological data or biomonitoring cohorts.

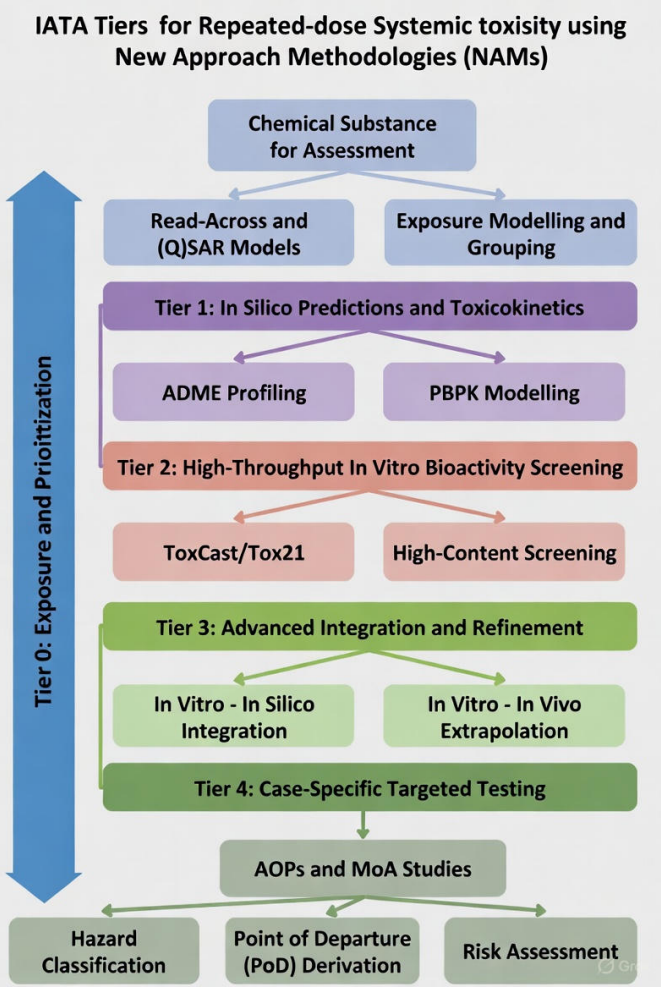
Feedback loop: Refine models with new NAM data or open literature.

4. Decision Rules and Outputs

Classification Example:

- ***Low concern***: No bioactivity below realistic exposures → unrestricted use/waiver.
- ***Medium:*** Bioactivity PoD supports safe level derivation.
- ***High****:* Strong signals → risk management or further data.

PoD for Risk Assessment: Conservative IVIVE-derived value with uncertainty factors.



5. Case Studies

Exemplar chemicals: e.g., pyrethroids (recent proof-of-concept), cosmetics ingredients, or pesticides from OECD chronic IATA cases.

Demonstrate concordance with historical 90-day data using NAM-only approaches.

6. Implementation Roadmap

Short-term: Pilot for specific classes (e.g., low-bioavailability chemicals).

Mid-term: Harmonize with OECD/ECHA/FDA.

Training: Regulator/industry workshops.

2. Outline for a Reproductive and Developmental Toxicity (DART) IATA Framework Using New Approach Methodologies (NAMs)

This outline proposes an exemplified Integrated Approach to Testing and Assessment (IATA) for reproductive and developmental toxicity (DART) endpoints (e.g., targeting fertility, embryotoxicity, teratogenicity, and developmental neurotoxicity as in OECD TG 414/416/421/422/443).

It draws from emerging 2025 frameworks, such as tiered NGRA approaches (e.g., ECETOC-inspired classification systems), physiologically-based kinetic (PBK) modeling integrated with in vitro bioactivity, and ongoing ICCVAM/OECD/EPA/FDA efforts to prioritize NAMs for DART endpoints.

The goal is a non-animal, human-relevant, tiered decision framework that provides points of departure (PoDs) for hazard identification, classification (e.g., Repr. 1A/1B/2), and risk assessment, while building regulatory confidence through weight-of-evidence (WoE).

1. Introduction and Scope

Purpose: To enable full replacement of in vivo DART studies (e.g., rodent two-generation or extended one-generation reproduction) with NAM batteries for industrial chemicals, pesticides, pharmaceuticals, and cosmetics.

Applicability Domain: Oral route initially; extension to inhalation/dermal. Focus on systemic DART effects (e.g., endocrine disruption, gametotoxicity, fetal malformations, neurodevelopmental impacts); excludes local effects.

Regulatory Alignment: Compatible with OECD IATA guidance, FDA Predictive Toxicology Roadmap (2025 updates emphasizing NAMs for developmental safety), ICCVAM DART strategies, and EU NGRA principles.

Decision Context: Hazard classification (low/medium/high concern per recent proposals), NOAEL/LOAEL estimation, or waiver for higher-tier animal testing.

2. Principles and Confidence-Building

Tiered Structure: Progressive information gathering to minimize testing while maximizing predictivity, focusing on sensitive windows (e.g., gestation, puberty).

Uncertainty Characterization: Applicability domain, reproducibility, mechanistic relevance (per OECD GD 34 updates, with emphasis on DART AOPs like thyroid disruption leading to cognitive deficits).

Validation Approach: Performance vs. historical animal/human data; use of open databases (e.g., ToxRefDB, ECHA DART dossiers, upcoming CAMERA beta mid-2025).

WoE Integration: Bayesian or structured expert judgment for combining data sources, incorporating exposure-led protectiveness (e.g., bioactivity-exposure ratios, BER).

3. Tiered IATA Workflow

Tier 0: Exposure and Prioritization (No Testing Needed for Low-Exposure Scenarios)

- Use Threshold of Toxicological Concern (TTC) adapted for DART (e.g., lower thresholds for reprotoxicants) or exposure-based waiving.
- Read-across/category approach from analogs (e.g., OECD QSAR Toolbox with DART-specific groupings).

Outcome: Low concern → no further testing; proceed otherwise.

Tier 1: In Silico Predictions and Toxicokinetics (TK)

- Physicochemical properties and ADME predictions (e.g., OPERA, VEGA tools, with placental transfer models).
- PBK/PBTK modeling for internal dosimetry (e.g., httk R package extended for pregnancy; reverse dosimetry to estimate maternal/fetal equivalent doses).
- (Q)SAR for DART endpoints (e.g., models for endocrine disruption via androgen/estrogen receptors; CATMoS-like for developmental NOAEL prediction).
- Bioavailability assessment (e.g., in vitro placental models for transplacental kinetics).

Outcome: Predict bioaccumulation/developmental potential; refine exposure; flag high-concern structures (e.g., teratogenic scaffolds).

Tier 2: High-Throughput In Vitro Bioactivity Screening

- ***Broad coverage assays:*** ToxCast/Tox21 battery (ER, AR, TR, etc.); high-content imaging for embryotoxicity.
- ***Targeted DART-specific assays:***
  - ***Reproductive***: H295R steroidogenesis (OECD TG 456); CALUX for ER/AR activity.
  - ***Developmental:*** Zebrafish embryo (FET, OECD TG 236); stem cell-based (e.g., devTOX quickPredict for teratogenicity; ReproTracker biomarkers).
  - ***Neurodevelopmental:*** iPSC-derived neural models (e.g., DNT battery per EFSA guidance).
  - Multi-organ chips (MPS) for maternal-fetal crosstalk (e.g., placenta-embryo systems).
- ***Repeated-exposure simulation:*** Dose fractionation over critical windows (e.g., organogenesis periods in vitro).
- ***PoD derivation:*** Benchmark concentration (BMC) from concentration-response curves, adjusted for cytotoxicity.

Outcome: Bioactivity PoD (e.g., lowest AC50 across assays); classify as inactive/low/medium/high potency.

SEE APPENDIX 1 Tier 2: High-Throughput In Vitro Bioactivity Screening

Tier 3: Advanced Integration and Refinement

IVIVE (In Vitro to In Vivo Extrapolation): Convert in vitro PoDs to external doses via PBK (accounting for gestational kinetics, fetal Css).

- ***AOP-informed WoE:*** Link bioactivity to key events (e.g., via AOP-Wiki for DART pathways like aromatase inhibition leading to fertility impairment).
- ***Omics integration:*** Transcriptomics/metabolomics for mode-of-action (e.g., TempO-Seq for DART biomarkers; pathway analysis for neurodevelopment).

- ***Uncertainty factors:*** Apply based on coverage gaps (e.g., 3-10 for extrapolation, higher for vulnerable populations).

Outcome: Human-relevant PoD (e.g., equivalent NOAEL in mg/kg/day); classification into low/medium/high concern.

Tier 4: Case-Specific Targeted Testing (If Needed)

Additional NAMs only if critical gaps (e.g., advanced organ-on-chip models, emerging omics, or enhanced read-across from larger analog sets); otherwise, rely on existing human epidemiological data or biomonitoring cohorts.

Feedback loop: Refine models with new NAM data or open literature.

4. Decision Rules and Outputs

Classification Example:
- ***Low concern***: No bioactivity below realistic exposures → unrestricted use/waiver.
- ***Medium*****:** Bioactivity PoD supports safe level derivation (e.g., BER > 100).
- ***High***: Strong signals → risk management or further data.

PoD for Risk Assessment: Conservative IVIVE-derived value with uncertainty factors, using BER = PoD / internal exposure.

5. **Case Studies**

Exemplar chemicals: e.g., phthalates (recent proof-of-concept for endocrine DART), pharmaceuticals like valproic acid, or pesticides from OECD DART IATA cases.

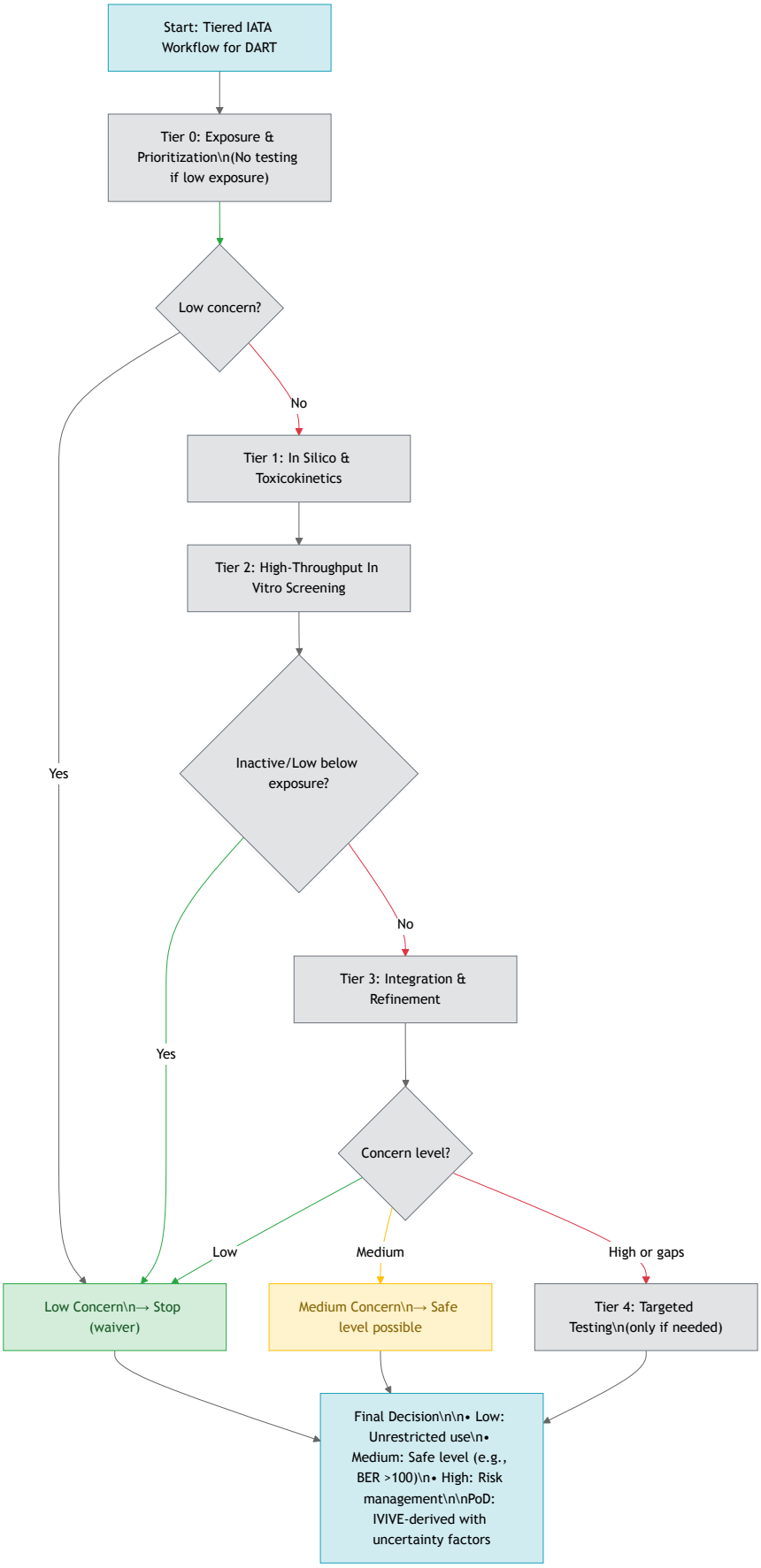
Demonstrate concordance with historical two-generation data using NAM-only approaches.

6. Implementation Roadmap

- ***Short-term:*** Pilot for specific classes (e.g., low-bioavailability reprotoxicants).
- ***Mid-term:*** Harmonize with OECD/ECHA/FDA.
- ***Training:*** Regulator/industry workshops.

See Appendix 2 : IVIVE in Tier 3 of the Tiered IATA Workflow for DART Assessment

Dart Workflow



3. Outline for a Carcinogenicity IATA Framework Using New Approach Methodologies (NAMs)

This outline proposes an exemplified Integrated Approach to Testing and Assessment (IATA) for carcinogenicity endpoints (e.g., targeting genotoxic and non-genotoxic mechanisms, including mutagenicity and tumor promotion as relevant to rodent bioassays and human relevance).

It draws from emerging 2025 frameworks, such as tiered NGRA approaches (e.g., ECETOC-inspired classification systems), physiologically-based kinetic (PBK) modeling integrated with in vitro bioactivity, AOP networks, and ongoing ICCVAM/OECD/EPA/FDA/PARC efforts to prioritize NAMs for carcinogenicity assessment (including ReCAAP and NGTxC IATA initiatives).

The goal is a non-animal, human-relevant, tiered decision framework that provides points of departure (PoDs) or weight-of-evidence conclusions for hazard identification, classification (e.g., Carc. 1A/1B/2), and risk assessment, while building regulatory confidence through weight-of-evidence (WoE).

1. Introduction and Scope

Purpose: To enable full replacement of in vivo carcinogenicity studies (e.g., rodent 2-year bioassay) with NAM batteries for industrial chemicals, pesticides, pharmaceuticals, and cosmetics.

Applicability Domain: Oral route initially; extension to inhalation/dermal. Focus on genotoxic and non-genotoxic mechanisms (e.g., DNA damage, cell proliferation, immunosuppression, endocrine modulation); addresses both mutagenic and non-mutagenic pathways.

Regulatory Alignment: Compatible with OECD IATA guidance (including NGTxC IATA and ReCAAP frameworks), FDA Predictive Toxicology Roadmap (2025 updates emphasizing NAMs for carcinogenicity), ICCVAM strategies, EFSA/PARC initiatives, and EU NGRA principles.

Decision Context: Hazard classification (low/medium/high concern per recent proposals), mutagenicity/carcinogenicity estimation, or waiver for higher-tier testing.

2. Principles and Confidence-Building

Tiered Structure: Progressive information gathering to minimize testing while maximizing predictivity, emphasizing mechanistic key events (hallmarks of cancer, AOPs).

Uncertainty Characterization: Applicability domain, reproducibility, mechanistic relevance (per OECD GD 34 updates, with emphasis on carcinogenicity AOP networks).

***Validation Approach*:** Performance vs. historical animal/human data; use of open databases (e.g., ToxRefDB, Carcinogenicity Potency Database, upcoming CAMERA beta mid-2025).

WoE Integration: Bayesian or structured expert judgment for combining data sources, incorporating exposure-led protectiveness (e.g., bioactivity-exposure ratios, BER).

3. Tiered IATA Workflow

Tier 0: Exposure and Prioritization (No Testing Needed for Low-Exposure Scenarios)

- Use Threshold of Toxicological Concern (TTC) adapted for carcinogens (e.g., lower thresholds for genotoxicants) or exposure-based waiving.

- Read-across/category approach from analogs (e.g., OECD QSAR Toolbox with carcinogenicity alerts).

Outcome*:* Low concern → no further testing; proceed otherwise.

Tier 1: In Silico Predictions and Toxicokinetics (TK)

- Physicochemical properties and ADME predictions (e.g., OPERA, VEGA tools).

- PBK/PBTK modeling for internal dosimetry (e.g., httk R package; reverse dosimetry to estimate chronic human equivalent doses).

- (Q)SAR and structural alerts for carcinogenicity (e.g., Toxtree, Derek Nexus, ISS models for mutagenicity/carcinogenicity).

Outcome**:** Predict genotoxic/non-genotoxic potential; refine exposure; flag high-concern structures (e.g., DNA-reactive alerts).

Tier 2: High-Throughput In Vitro Bioactivity Screening

Broad coverage assays**:** ToxCast/Tox21 battery (including genotoxicity-relevant endpoints like p53, ATAD5).

Targeted carcinogenicity-specific assays:

- ***Genotoxicity***: Ames (OECD TG 471 bacterial reverse mutation), in vitro micronucleus (OECD TG 487), MultiFlow/Micronucleus assays, ToxTracker reporter assay.

- ***Non-genotoxic:*** Cell transformation assays (e.g., Bhas 42, SHE CTA per OECD GD), hallmarks of cancer assays (proliferation, apoptosis resistance, immortalization).

- ***Mechanistic:*** High-content imaging for key events (e.g., oxidative stress, receptor-mediated like AhR/CAR/PXR).

- ***Omics integration:*** Transcriptomics for carcinogen signatures (e.g., TGx-DDI biomarker for DNA damage).

- ***PoD derivation:*** Benchmark concentration (BMC) from concentration-response curves.

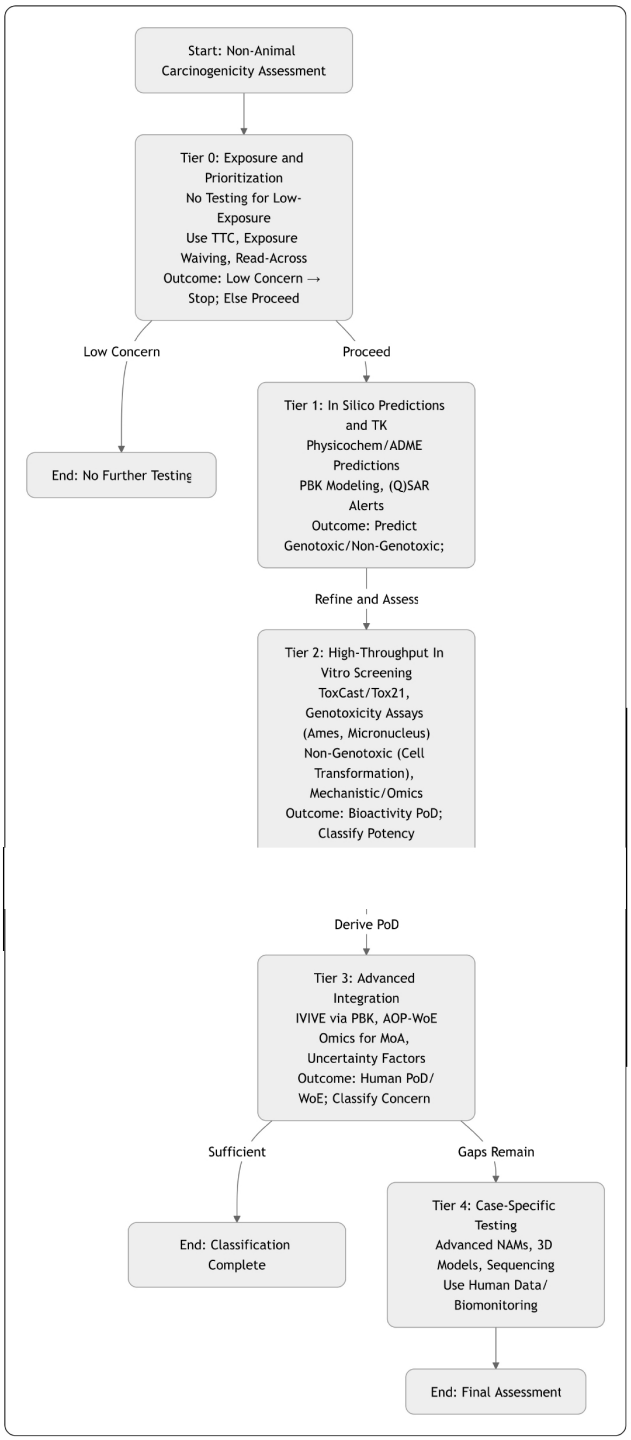
Outcome**:** Bioactivity PoD (e.g., lowest AC50); classify as genotoxic/non-genotoxic, low/medium/high potency.

Tier 3: Advanced Integration and Refinement

- ***IVIVE (In Vitro to In Vivo Extrapolation):*** Convert in vitro PoDs to external doses via PBK (accounting for chronic exposure).

- ***AOP-informed WoE:*** Link bioactivity to key events/hallmarks (e.g., via AOP-Wiki networks for genotoxic and non-genotoxic pathways).

- ***Omics integration:*** Transcriptomics/metabolomics for mode-of-action (e.g., TempO-Seq for carcinogenicity biomarkers).



-***Uncertainty factors:*** Apply based on coverage gaps (e.g., 3-10 for extrapolation, higher for mechanistic uncertainty).

Outcome**:** Human-relevant PoD or WoE conclusion; classification into low/medium/ high concern.

Tier 4: Case-Specific Targeted Testing (If Needed)

Additional NAMs only if critical gaps (e.g., advanced 3D models, error-corrected sequencing for mutational signatures, enhanced AOP mapping, or larger analog read-across); otherwise, rely on existing human epidemiological data or biomonitoring cohorts.

Feedback loop**:** Refine models with new NAM data or open literature.

4. Decision Rules and Outputs

Classification Example (Adapted from 2025 proposals):

- ***Low concern***: No bioactivity/mechanistic signals below realistic exposures → unrestricted use/waiver.
- ***Medium:*** Bioactivity PoD supports safe level derivation (e.g., BER > 1000 for genotoxicants).
- ***High*****:** Strong genotoxic/non-genotoxic signals → risk management or further data.

PoD for Risk Assessment**:** Conservative IVIVE-derived value with uncertainty factors, using BER = PoD / internal exposure.

5. Case Studies

Exemplar chemicals**:** e.g., genotoxicants like aflatoxin B1, non-genotoxic like phthalates or PPAR agonists, or agrochemicals from OECD/ReCAAP IATA cases.

Demonstrate concordance with historical bioassay data using NAM-only approaches.

6. Implementation Roadmap

- ***Short-term*****:** Pilot for specific classes (e.g., data-rich genotoxicants or low-exposure compounds).
- ***Mid-term*****:** Harmonize with OECD/ECHA/FDA/ICCVAM/PARC.
- ***Training*****:** Regulator/industry workshops.

4. Outline for a Inhalation Toxicity (Respiratory Tract) IATA Framework Using New Approach Methodologies (NAMs)

This outline proposes an exemplified Integrated Approach to Testing and Assessment (IATA) for inhalation toxicity endpoints focusing on the respiratory tract (e.g., acute irritation, portal-of-entry effects, repeated-dose local effects like inflammation or fibrosis, as relevant to OECD TG 403/412/413 and point-of-contact toxicity).

It draws from emerging 2025 frameworks, such as tiered NGRA approaches (e.g., consumer goods-inspired classification systems), dosimetry modeling (MPPD/ICRP/CFD) integrated with in vitro bioactivity at air-liquid interface (ALI), AOP networks for respiratory irritation/sensitization, and ongoing ICCVAM/OECD/EPA/PARC efforts to prioritize NAMs for inhalation endpoints (including OECD No. 367 case study and NGRA inhalation toolboxes).

The goal is a non-animal, human-relevant, tiered decision framework that provides points of departure (PoDs) for local respiratory hazard identification, classification (e.g., acute inh. tox., STOT-SE/RE), and risk assessment, while building regulatory confidence through weight-of-evidence (WoE).

1. Introduction and Scope

***Purpose*:** To enable full replacement of in vivo inhalation studies (e.g., rodent acute/subchronic nose-only or whole-body exposure) with NAM batteries for industrial chemicals, pesticides, pharmaceuticals, cosmetics, and consumer products.

***Applicability Domain*:** Inhalation route (gases, vapors, aerosols/particles). Focus on portal-of-entry effects in the respiratory tract (upper: nasal/bronchial; lower: alveolar; irritation, inflammation, cytotoxicity, fibrosis); systemic effects may be addressed separately or via integration.

***Regulatory Alignment*:** Compatible with OECD IATA guidance (e.g., No. 367 for point-of-contact refinement), FDA Predictive Toxicology Roadmap (2025 updates emphasizing NAMs for inhalation safety), ICCVAM/EPA strategies, EFSA/PARC initiatives, and EU NGRA principles.

***Decision Context*:** Hazard classification (low/medium/high concern per recent proposals), NOAEC/LOAEC estimation, or waiver for higher-tier testing.

2. Principles and Confidence-Building

***Tiered Structure*:** Progressive information gathering to minimize testing while maximizing predictivity, emphasizing regional dosimetry and sensitive exposure windows (acute vs. repeated).

Uncertainty Characterization: Applicability domain, reproducibility, mechanistic relevance (per OECD GD 34 updates, with emphasis on respiratory AOPs like protein alkylation leading to irritation or oxidative stress to fibrosis).

Validation Approach: Performance vs. historical animal/human data; use of open databases (e.g., ToxRefDB, ECHA inhalation dossiers, upcoming CAMERA beta mid-2025).

WoE Integration: Bayesian or structured expert judgment for combining data sources, incorporating exposure-led protectiveness (e.g., bioactivity-exposure ratios adjusted for deposited dose, BER).

3. Tiered IATA Workflow

Tier 0: Exposure and Prioritization (No Testing Needed for Low-Exposure Scenarios)

- Use Threshold of Toxicological Concern (TTC) adapted for inhalation (e.g., lower for irritants) or exposure-based waiving.

- Read-across/category approach from analogs (e.g., OECD QSAR Toolbox with respiratory alerts).

Outcome: Low concern → no further testing; proceed otherwise.

Tier 1: In Silico Predictions and Dosimetry (TK for Inhalation)

- ***Physicochemical properties:*** (volatility, particle size, reactivity) and predictions (e.g., OPERA, VEGA tools).

- ***Inhalation dosimetry modeling:*** MPPD for regional deposition; CFD for upper airways; ICRP clearance models.

- ***(Q)SAR and structural alerts for respiratory toxicity*** (e.g., irritant/corrosive alerts, reactive mechanisms).

Outcome: Predict deposited dose and regional targeting; refine exposure; flag high-concern structures (e.g., reactive gases).

Tier 2: High-Throughput In Vitro Bioactivity Screening at ALI

- ***Broad coverage assays:*** ToxCast/Tox21 subsets relevant to respiratory mechanisms (e.g., oxidative stress, inflammation pathways).

- ***Targeted respiratory-specific assays:***

  - ***Upper tract*****:** MucilAir™ or similar nasal/bronchial ALI models (e.g., cytotoxicity, TEER, cytokine release).

  - ***Lower tract:*** EpiAlveolar™ or small airway/alveolar ALI models (e.g., barrier integrity, inflammation markers).

  - ***Advanced:*** Multi-cell/3D organotypic models for fibrosis or sensitization.

- ***Exposure simulation*****:** Aerosol/vapor delivery systems at ALI; repeated exposures for subchronic relevance.

- ***PoD derivation***: Benchmark concentration (BMC) from concentration-response curves, adjusted for cytotoxicity and deposited dose.

Outcome**:** Bioactivity PoD (e.g., lowest effective concentration); classify as inactive/low/medium/high potency by region.

Tier 3: Advanced Integration and Refinement

IVIVE (In Vitro to In Vivo Extrapolation): Convert in vitro PoDs to external/deposited doses via dosimetry models (accounting for clearance, repeated exposure).

AOP-informed WoE: Link bioactivity to key events (e.g., via AOP-Wiki for respiratory irritation or sensitization pathways).

Omics integration**:** Transcriptomics/metabolomics for mode-of-action (e.g., TempO-Seq for irritation/fibrosis biomarkers).

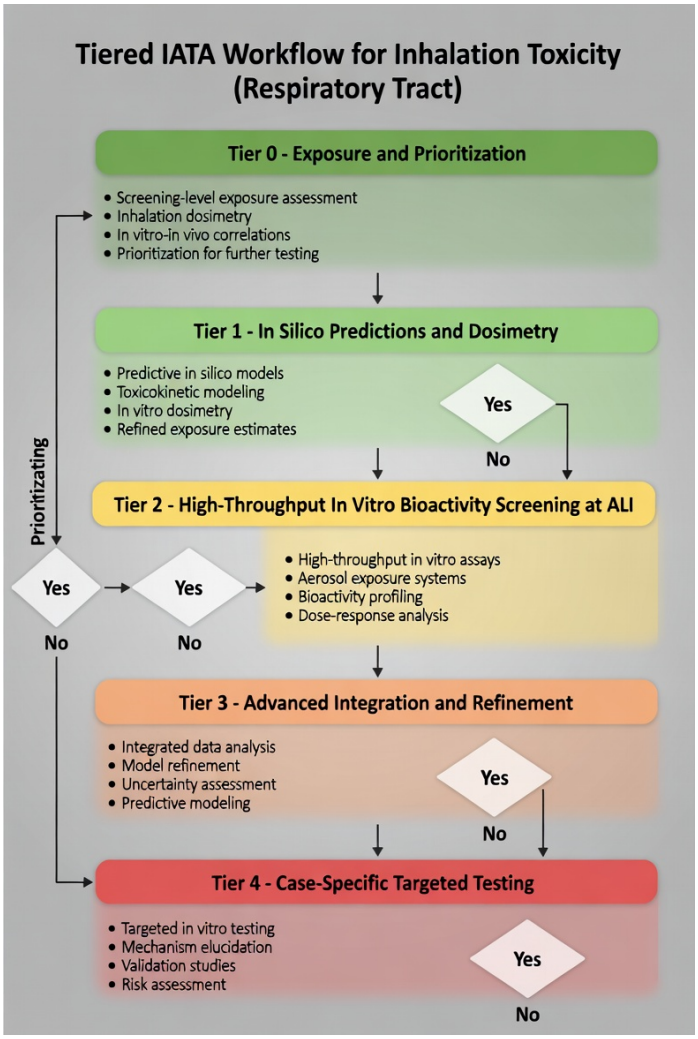
Uncertainty factors: Apply based on coverage gaps (e.g., 3-10 for regional extrapolation, dosimetry uncertainty).

Outcome: Human-relevant PoD (e.g., equivalent NOAEC in mg/m³ or µg/cm² deposited); classification into low/medium/high concern.

Tier 4: Case-Specific Targeted Testing (If Needed)

Additional NAMs only if critical gaps (e.g., advanced 3D lung-on-chip models, enhanced dosimetry, emerging omics, or larger analog read-across); otherwise, rely on existing human epidemiological data or biomonitoring cohorts.

***Feedback loop*:** Refine models with new NAM data or open literature.



4. Decision Rules and Outputs

Classification Example:

- ***Low concern*:** No bioactivity below realistic deposited exposures → unrestricted use/waiver.

- ***Medium*:** Bioactivity PoD supports safe level derivation (e.g., BER > 100).

- ***High*:** Strong regional signals → risk management or further data.

PoD for Risk Assessment: Conservative IVIVE-derived value with uncertainty factors, using BER = PoD / deposited internal exposure.

5. Case Studies

Exemplar chemicals**:** e.g., irritants like chlorothalonil (OECD No. 367 proof-of-concept), cleaning product ingredients, or aerosols from NGRA inhalation cases.

Demonstrate concordance with historical inhalation data using NAM-only approaches.

6. Implementation Roadmap

- ***Short-term*****:** Pilot for specific classes (e.g., low-volatility consumer ingredients or known irritants).
- ***Mid-term*****:** Harmonize with OECD/ECHA/FDA/ICCVAM/PARC.
- ***Training*****:** Regulator/industry workshops.

5. Outline for a Acute Systemic Toxicity (Oral/Inhalation, including Lethality Endpoints) IATA Framework Using New Approach Methodologies (NAMs)

This outline proposes an exemplified Integrated Approach to Testing and Assessment (IATA) for acute systemic toxicity endpoints (oral and inhalation routes, including lethal outcomes such as LD50/LC50, as relevant to OECD TG 423/425/403/436 and GHS classification categories).

It draws from emerging 2025 frameworks, such as tiered NGRA approaches, in silico consensus modeling (e.g., CATMoS/CoMPAIT), in vitro bioactivity profiling, dosimetry for inhalation, AOP networks for acute mechanisms, and ongoing ICCVAM/OECD/EPA/PARC efforts to prioritize NAMs for acute lethality prediction.

The goal is a non-animal, human-relevant, tiered decision framework that provides points of departure (PoDs) or probabilistic estimates for hazard identification, classification (e.g., Acute Tox. 1-5), and risk assessment, while building regulatory confidence through weight-of-evidence (WoE).

1. Introduction and Scope

Purpose: To enable full replacement of in vivo acute systemic toxicity studies (e.g., rodent oral LD50 or inhalation LC50) with NAM batteries for industrial chemicals, pesticides, pharmaceuticals, cosmetics, and consumer products.

Applicability Domain: Oral and inhalation routes (gases, vapors, aerosols/particles); focus on systemic effects post-absorption/deposition, including lethality; local respiratory effects may integrate separately.

Regulatory Alignment: Compatible with OECD IATA guidance (e.g., GD 237 on waiving/bridging), FDA Predictive Toxicology Roadmap (2025 updates emphasizing NAMs for acute safety), ICCVAM Acute Toxicity strategies (CATMoS/CoMPAIT), and EU NGRA principles.

Decision Context: Hazard classification (low/medium/high concern or GHS categories), LD50/LC50 estimation, or waiver for higher-tier testing.

2. Principles and Confidence-Building

Tiered Structure: Progressive information gathering to minimize testing while maximizing predictivity, emphasizing mechanistic key events (e.g., basal cytotoxicity, oxidative stress, organ-specific disruption).

Uncertainty Characterization: Applicability domain, reproducibility, mechanistic relevance (per OECD GD 34 updates, with emphasis on acute AOPs).

Validation Approach: Performance vs. historical animal/human data; use of open databases (e.g., curated rat LD50 datasets from ICCVAM, upcoming CAMERA beta mid-2025).

WoE Integration: Bayesian or structured expert judgment for combining data sources, incorporating exposure-led protectiveness (e.g., bioactivity-exposure ratios, BER).

3. Tiered IATA Workflow

Tier 0: Exposure and Prioritization (No Testing Needed for Low-Exposure Scenarios)

- Use Threshold of Toxicological Concern (TTC) adapted for acute (oral/inhalation) or exposure-based waiving.
- Read-across/category approach from analogs (e.g., OECD QSAR Toolbox with acute alerts).

Outcome: Low concern → no further testing; proceed otherwise.

Tier 1: In Silico Predictions and Toxicokinetics/Dosimetry

- Physicochemical properties and ADME/reactivity predictions (e.g., OPERA, VEGA tools).
- PBK modeling for oral; inhalation dosimetry (e.g., MPPD for deposited dose, CFD for regional targeting).
- (Q)SAR and consensus models for acute endpoints (e.g., CATMoS for oral LD50/GHS; emerging CoMPAIT for inhalation LC50).
- Bioavailability/permeability assessment (e.g., in vitro Caco-2 for oral).

Outcome: Predict acute potential; refine exposure/dose; flag high-concern structures (e.g., reactive alerts).

Tier 2: High-Throughput In Vitro Bioactivity Screening

- ***Broad coverage assays:*** ToxCast/Tox21 battery (cytotoxicity, mitochondrial disruption, etc.); 3T3-NRU basal cytotoxicity (for oral starting dose/LD50 prediction).
- ***Targeted acute-specific assays:***
  - ***General:*** High-content imaging for key events (oxidative stress, apoptosis).
  - ***Organ-specific***: HepaRG/hepatocytes, neuronal models, or multi-organ chips for systemic targets.
  - ***Inhalation***: ALI respiratory models (e.g., MucilAir/EpiAlveolar) for absorbed fraction/systemic spillover.
  - ***Exposure simulation:*** Acute dosing; aerosol delivery for inhalation relevance.
  - ***PoD derivation:*** Benchmark concentration (BMC) or IC50 from concentration-response, adjusted for cytotoxicity.

Outcome: Bioactivity PoD (e.g., lowest AC50); classify as inactive/low/medium/high potency.

Tier 3: Advanced Integration and Refinement

- ***IVIVE (In Vitro to In Vivo Extrapolation):*** Convert in vitro PoDs to external/deposited doses via PBK/dosimetry (accounting for acute kinetics).
- ***AOP-informed WoE:*** Link bioactivity to key events (e.g., via AOP-Wiki for acute systemic pathways).
- ***Omics integration:*** Transcriptomics for acute signatures (e.g., TempO-Seq biomarkers).

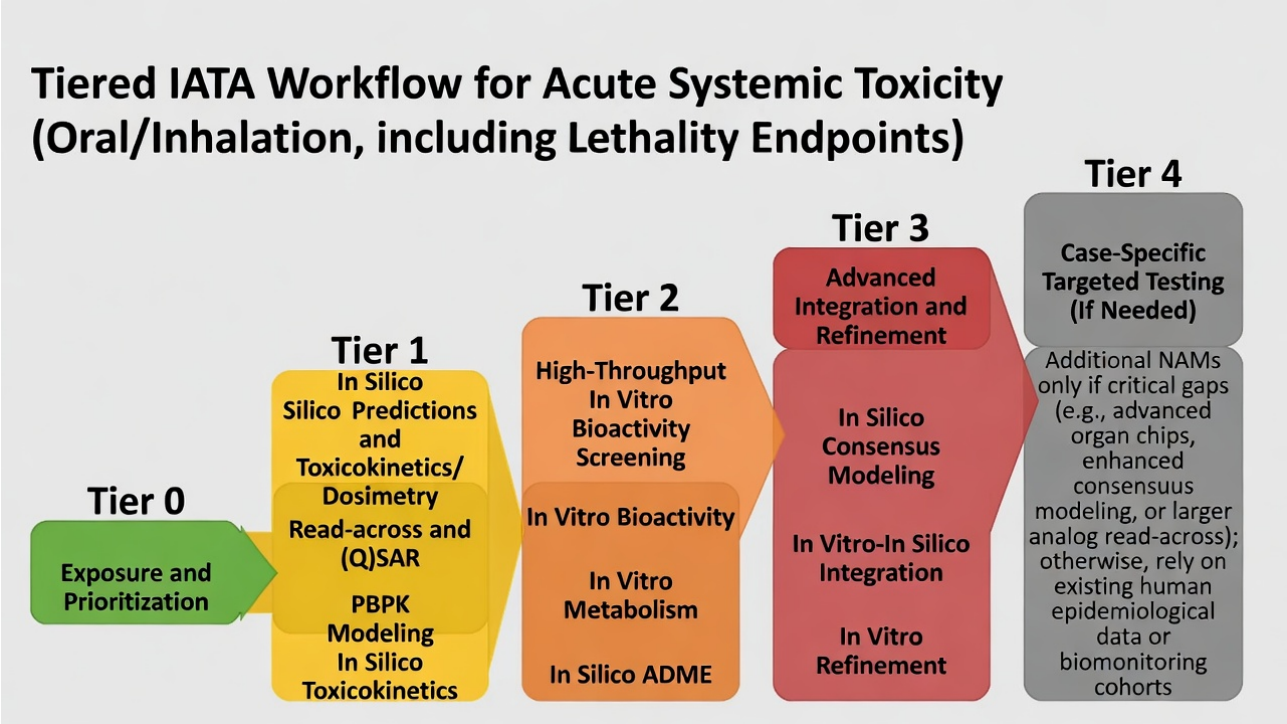
Uncertainty factors: Apply based on coverage gaps (e.g., 3-10 for extrapolation).

Outcome: Human-relevant PoD (e.g., equivalent LD50/LC50 or category); classification into low/medium/high concern.

Tier 4: Case-Specific Targeted Testing (If Needed)

Additional NAMs only if critical gaps (e.g., advanced organ chips, enhanced consensus modeling, or larger analog read-across); otherwise, rely on existing human epidemiological data or biomonitoring cohorts.

Feedback loop: Refine models with new NAM data or open literature.



4. Decision Rules and Outputs

Classification Example:

- ***Low concern:*** No bioactivity below realistic exposures → unrestricted use/waiver (e.g., >2000 mg/kg oral or non-classified inhalation).
- ***Medium***: Bioactivity PoD supports safe level derivation.
- ***High:*** Strong signals → risk management or further data.

PoD for Risk Assessment: Conservative IVIVE-derived value with uncertainty factors, using BER = PoD / internal exposure.

5. Case Studies

Exemplar chemicals: e.g., pharmaceuticals/cosmetics with low acute concern, pesticides/irritants from ICCVAM/CATMoS/CoMPAIT cases, or volatiles for inhalation.

Demonstrate concordance with historical acute data using NAM-only approaches.

6. Implementation Roadmap

- ***Short-term:*** Pilot for specific classes (e.g., low-volatility or data-rich compounds).
- ***Mid-term:*** Harmonize with OECD/ECHA/FDA/ICCVAM/PARC.
- ***Training:*** Regulator/industry workshops.

6. Outline for a Neurotoxicity and Behavioral Endpoints IATA Framework Using New Approach Methodologies (NAMs)

This outline proposes an exemplified Integrated Approach to Testing and Assessment (IATA) for neurotoxicity endpoints, encompassing both developmental neurotoxicity (DNT) and adult neurotoxicity (ANT), including behavioral outcomes (e.g., targeting learning/memory impairment, motor deficits, sensory alterations as relevant to OECD TG 424/426/443 and functional observational batteries).

It draws from emerging 2025 frameworks, such as tiered NGRA approaches, AOP-informed IATA (e.g., OECD DNT-IVB guidance, EFSA/PARC initiatives), physiologically-based kinetic (PBK) modeling integrated with in vitro bioactivity, and ongoing ICCVAM/OECD/EPA/EFSA efforts to prioritize NAMs for neurotoxicity (including DNT in vitro battery expansions and ANT mechanistic assays).

The goal is a non-animal, human-relevant, tiered decision framework that provides points of departure (PoDs) for hazard identification, classification (e.g., STOT-RE with neuro specificity), and risk assessment, while building regulatory confidence through weight-of-evidence (WoE).

1. Introduction and Scope

Purpose: To enable full replacement of in vivo neurotoxicity studies (e.g., rodent DNT cohorts or adult functional tests) with NAM batteries for industrial chemicals, pesticides, pharmaceuticals, and cosmetics.

Applicability Domain: Oral route initially; extension to inhalation/dermal. Focus on developmental (e.g., neural proliferation, migration, synaptogenesis) and adult (e.g., neurodegeneration, network function) effects, including behavioral endpoints (motor activity, cognition, sensory function).

Regulatory Alignment: Compatible with OECD IATA guidance (e.g., DNT-IVB recommendations, AOP networks), FDA Predictive Toxicology Roadmap (2025 updates emphasizing NAMs for neuro safety), ICCVAM/EFSA/PARC strategies, and EU NGRA principles.

Decision Context: Hazard classification (low/medium/high concern per recent proposals), NOAEL estimation, or waiver for higher-tier testing.

2. Principles and Confidence-Building

Tiered Structure: Progressive information gathering to minimize testing while maximizing predictivity, focusing on sensitive windows (gestation for DNT, chronic for ANT) and behavioral surrogates.

Uncertainty Characterization: Applicability domain, reproducibility, mechanistic relevance (per OECD GD 34 updates, with emphasis on neurotoxicity AOP networks).

Validation Approach: Performance vs. historical animal/human data; use of open databases (e.g., ToxRefDB, EFSA DNT compendium, upcoming CAMERA beta mid-2025).

WoE Integration: Bayesian or structured expert judgment for combining data sources, incorporating exposure-led protectiveness (e.g., bioactivity-exposure ratios, BER).

3. Tiered IATA Workflow

Tier 0: Exposure and Prioritization (No Testing Needed for Low-Exposure Scenarios)

- Use Threshold of Toxicological Concern (TTC) adapted for neurotoxicants or exposure-based waiving.
- Read-across/category approach from analogs (e.g., OECD QSAR Toolbox with neuro alerts).

Outcome: Low concern → no further testing; proceed otherwise.

Tier 1: In Silico Predictions and Toxicokinetics (TK)

- Physicochemical properties and ADME predictions (e.g., OPERA, VEGA tools, with blood-brain barrier models).
- PBK/PBTK modeling for internal dosimetry (e.g., httk R package; reverse dosimetry to estimate brain/target equivalent doses).
- (Q)SAR for neurotoxicity endpoints (e.g., models for receptor binding, mitochondrial disruption; CATMoS-like for behavioral NOAEL prediction).

Outcome: Predict brain penetration/neuro potential; refine exposure; flag high-concern structures (e.g., VGSC inhibitors).

Tier 2: High-Throughput In Vitro Bioactivity Screening

Broad coverage assays: ToxCast/Tox21 battery (neuronal receptors, mitochondrial, oxidative stress).

Targeted neuro-specific assays:

- DNT: OECD-aligned DNT-IVB (e.g., neural progenitor proliferation, migration, neurite outgrowth, synaptogenesis in hiPSC models; network formation on MEA).
- ANT: Mature neuronal models (e.g., hiPSC cortical/spinal networks, microelectrode arrays for functional activity).
- Behavioral surrogates: High-content imaging for key events (e.g., synaptic plasticity); omics for cognitive/motor signatures.

Alternative models: Zebrafish embryo/larval behavior (non-protected stages, e.g., locomotor, photomotor response).

Exposure simulation: Acute/repeated dosing; developmental windows in vitro.

PoD derivation: Benchmark concentration (BMC) from concentration-response curves.

Outcome: Bioactivity PoD (e.g., lowest AC50); classify as inactive/low/medium/high potency.

Tier 3: Advanced Integration and Refinement

- ***IVIVE (In Vitro to In Vivo Extrapolation):*** Convert in vitro PoDs to external doses via PBK (accounting for brain kinetics).
- ***AOP-informed WoE:*** Link bioactivity to key events (e.g., via AOP-Wiki networks for VGSC inhibition to cognitive impairment or mitochondrial complex I to parkinsonian deficits).
- ***Omics integration:*** Transcriptomics/metabolomics for mode-of-action (e.g., TempO-Seq for neuro biomarkers).

Uncertainty factors: Apply based on coverage gaps (e.g., 3-10 for extrapolation, higher for behavioral prediction).

Outcome: Human-relevant PoD (e.g., equivalent NOAEL in mg/kg/day); classification into low/medium/high concern.

Tier 4: Case-Specific Targeted Testing (If Needed)

Additional NAMs only if critical gaps (e.g., advanced brain organoids/spheroids, enhanced MEA for seizure liability, emerging omics, or larger analog read-across); otherwise, rely on existing human epidemiological data or biomonitoring cohorts.

Feedback loop: Refine models with new NAM data or open literature.

4. Decision Rules and Outputs

Classification Example:

- ***Low concern:*** No bioactivity below realistic exposures → unrestricted use/waiver.

- ***Medium***: Bioactivity PoD supports safe level derivation (e.g., BER > 100).

- ***High:*** Strong signals (e.g., DNT-IVB hits or functional disruption) → risk management or further data.

PoD for Risk Assessment: Conservative IVIVE-derived value with uncertainty factors, using BER = PoD / internal brain exposure.

5. Case Studies

Exemplar chemicals: e.g., pyrethroids/deltamethrin (VGSC AOP proof-of-concept), rotenone (parkinsonian models), or pesticides from OECD/EFSA DNT IATA cases.

Demonstrate concordance with historical neuro data using NAM-only approaches.

6. Implementation Roadmap

- ***Short-term:*** Pilot for specific classes (e.g., known mitochondrial disruptors or low-BBB penetrants).
- ***Mid-term:*** Harmonize with OECD/EFSA/FDA/PARC (e.g., DNT-IVB v2.0 expansions).
- ***Training:*** Regulator/industry workshops.

7.Outline for an Immunotoxicity (beyond Skin Sensitization) IATA Framework Using New Approach Methodologies (NAMs)

This outline proposes an exemplified Integrated Approach to Testing and Assessment (IATA) for immunotoxicity endpoints beyond skin sensitisation (e.g., systemic immunosuppression, immunostimulation, autoimmunity, and hypersensitivity, targeting effects as relevant to OECD TG 443 extended one-generation and human immune-related adverse outcomes).

It draws from emerging 2025 frameworks, such as tiered NGRA approaches, AOP-informed IATA (e.g., OECD immunosuppression AOP network, PARC immunotoxicity initiatives), physiologically-based kinetic (PBK) modeling integrated with in vitro bioactivity, and ongoing ICCVAM/OECD/EPA/EFSA efforts to prioritize NAMs for systemic immune effects (including myeloid/lymphoid suppression and cytokine dysregulation).

The goal is a non-animal, human-relevant, tiered decision framework that provides points of departure (PoDs) for hazard identification, classification (e.g., specific immune-related concerns), and risk assessment, while building regulatory confidence through weight-of-evidence (WoE).

1. Introduction and Scope

Purpose: To enable full replacement of in vivo immunotoxicity studies (e.g., rodent T-dependent antibody response, immunophenotyping, or functional assays) with NAM batteries for industrial chemicals, pesticides, pharmaceuticals, and cosmetics.

Applicability Domain: Oral route initially; extension to inhalation/dermal. Focus on systemic immunotoxicity (immunosuppression, immunostimulation, autoimmunity, non-respiratory hypersensitivity); excludes skin sensitisation (addressed separately via OECD-defined key events).

Regulatory Alignment: Compatible with OECD IATA guidance (e.g., GD on immunosuppression AOPs), FDA Predictive Toxicology Roadmap (2025 updates emphasizing NAMs for immune safety), ICCVAM/EFSA/PARC strategies, and EU NGRA principles.

Decision Context: Hazard classification (low/medium/high concern per recent proposals), immune PoD estimation, or waiver for higher-tier testing.

2. Principles and Confidence-Building

Tiered Structure: Progressive information gathering to minimize testing while maximizing predictivity, focusing on key immune cell populations and functions.

Uncertainty Characterization: Applicability domain, reproducibility, mechanistic relevance (per OECD GD 34 updates, with emphasis on immunotoxicity AOP networks).

Validation Approach: Performance vs. historical animal/human data; use of open databases (e.g., ToxRefDB immune endpoints, ECHA dossiers, upcoming CAMERA beta mid-2025).

WoE Integration: Bayesian or structured expert judgment for combining data sources, incorporating exposure-led protectiveness (e.g., bioactivity-exposure ratios, BER).

3. Tiered IATA Workflow

Tier 0: Exposure and Prioritization (No Testing Needed for Low-Exposure Scenarios)

- Use Threshold of Toxicological Concern (TTC) adapted for immunotoxicants or exposure-based waiving.
- Read-across/category approach from analogs (e.g., OECD QSAR Toolbox with immune alerts).

Outcome: Low concern → no further testing; proceed otherwise.

Tier 1: In Silico Predictions and Toxicokinetics (TK)

- Physicochemical properties and ADME predictions (e.g., OPERA, VEGA tools).
- PBK/PBTK modeling for internal dosimetry (e.g., httk R package; reverse dosimetry to estimate immune compartment doses).
- (Q)SAR for immunotoxicity endpoints (e.g., structural alerts for immunosuppression, machine learning models for cytokine modulation).

Outcome: Predict immune targeting potential; refine exposure; flag high-concern structures (e.g., known immunosuppressants).

Tier 2: High-Throughput In Vitro Bioactivity Screening

Broad coverage assays: ToxCast/Tox21 battery (relevant immune pathways, e.g., NF-κB, STAT, glucocorticoid receptor).

Targeted immune-specific assays:

- ***Lymphoid***: Human PBMC or co-culture models (T-cell proliferation, cytokine release, e.g., IL-2/IFN-γ suppression).

- ***Myeloid:*** THP-1 or dendritic cell models (phagocytosis, maturation markers, inflammasome activation).

- ***Functional:*** T-cell dependent antibody response surrogates (e.g., co-culture with B-cells); immunosuppression panels (e.g., ImmunoSearch, BioSeek).

- ***Autoimmunity/hypersensitivity:*** High-content assays for T-cell activation markers or autoantibody induction surrogates.

Exposure simulation: Repeated dosing to capture chronic effects; 3D lymphoid organoids for context.

PoD derivation: Benchmark concentration (BMC) from concentration-response curves, adjusted for cytotoxicity.

Outcome: Bioactivity PoD (e.g., lowest effective concentration); classify as inactive/low/medium/high potency.

Tier 3: Advanced Integration and Refinement

- ***IVIVE (In Vitro to In Vivo Extrapolation):*** Convert in vitro PoDs to external doses via PBK (accounting for immune cell kinetics).
- ***AOP-informed WoE:*** Link bioactivity to key events (e.g., via AOP-Wiki networks for T-cell suppression to immunosuppression or NLRP3 activation to autoimmunity).
- ***Omics integration:*** Transcriptomics/metabolomics for mode-of-action (e.g., TempO-Seq for immune gene signatures).

Uncertainty factors: Apply based on coverage gaps (e.g., 3-10 for extrapolation, higher for adaptive immunity complexity).

Outcome: Human-relevant PoD (e.g., equivalent immune NOAEL in mg/kg/day); classification into low/medium/high concern.

Tier 4: Case-Specific Targeted Testing (If Needed)

Additional NAMs only if critical gaps (e.g., advanced 3D lymph node/immune organoids, enhanced functional co-cultures, emerging omics, or larger analog read-across); otherwise, rely on existing human epidemiological data or biomonitoring cohorts.

Feedback loop: Refine models with new NAM data or open literature.

4. Decision Rules and Outputs

Classification Example:

- ***Low concern***: No bioactivity below realistic exposures → unrestricted use/waiver.

- ***Medium***: Bioactivity PoD supports safe level derivation (e.g., BER > 100).

- ***High***: Strong signals (e.g., consistent cytokine suppression or T-cell impairment) → risk management or further data.

PoD for Risk Assessment: Conservative IVIVE-derived value with uncertainty factors, using BER = PoD / internal exposure.

5. Case Studies

Exemplar chemicals: e.g., cyclosporine/tacrolimus (known immunosuppressants), azathioprine, or pesticides/pharmaceuticals from OECD/PARC immunotoxicity cases.

Demonstrate concordance with historical immune data using NAM-only approaches.

6. Implementation Roadmap

- ***Short-term:*** Pilot for specific classes (e.g., known cytokine modulators or low-immune penetrants).
- ***Mid-term:*** Harmonize with OECD/EFSA/FDA/PARC (e.g., expanding immunosuppression AOP coverage).
- ***Training:*** Regulator/industry workshops.

8.Outline for an Ecotoxicology (Environmental Impact) IATA Framework Using New Approach Methodologies (NAMs)

This outline proposes an exemplified Integrated Approach to Testing and Assessment (IATA) for ecotoxicology endpoints (e.g., acute and chronic toxicity to aquatic and terrestrial organisms, sediment effects, bioaccumulation, and secondary poisoning, as relevant to OECD TG 201/202/203/210/211/215/218/225/235 and REACH environmental hazard classification).

It draws from emerging 2025 frameworks, such as tiered sequence-based approaches (e.g., SEQ-IATA, ECHA/PARC initiatives), in silico profiling (e.g., ECOSAR, VEGA), species sensitivity distribution (SSD) modeling, bioaccumulation predictions (e.g., BCFBAF), and ongoing OECD/ECHA/EFSA/PARC efforts to prioritize NAMs for environmental risk assessment (including fish embryo toxicity alternatives and mechanistic read-across).

The goal is a non-animal, mechanism-informed, tiered decision framework that provides predicted no-effect concentrations (PNEC) or hazard classifications for environmental compartments, supporting PBT/vPvB assessment and risk characterization, while building regulatory confidence through weight-of-evidence (WoE).

1. Introduction and Scope

Purpose: To enable full replacement of in vivo ecotoxicology studies (e.g., fish acute/chronic, daphnia immobilization/reproduction, algal growth) with NAM batteries for industrial chemicals, pesticides, pharmaceuticals, and biocides.

Applicability Domain: Freshwater, marine, sediment, and terrestrial compartments; focus on algae, invertebrates, fish, birds/mammals (via secondary poisoning), and microbial function. Includes bioaccumulation and persistence where relevant.

Regulatory Alignment: Compatible with OECD IATA guidance (e.g., GD 356 on fish toxicity alternatives), ECHA REACH guidance (R.7b/R.10/R.11 updates emphasizing NAMs), EFSA aquatic guidance (2025 revisions), EPA ECOTOX strategies, and EU NGRA principles for environmental endpoints.

Decision Context: Environmental hazard classification (e.g., Aquatic Acute/Chronic 1-3, PBT/vPvB), PNEC derivation, or waiver for higher-tier testing.

2. Principles and Confidence-Building

Tiered Structure: Progressive information gathering to minimize testing while maximizing predictivity, emphasizing mode-of-action (MoA) and cross-species extrapolation.

Uncertainty Characterization: Applicability domain, reproducibility, mechanistic relevance (per OECD GD 34 updates, with emphasis on ecotoxicity AOPs and QIVIVE).

Validation Approach: Performance vs. historical experimental data; use of open databases (e.g., ECOTOX, ECHA dossiers, REACH registrations, EnviroTox).

WoE Integration: Bayesian or structured expert judgment for combining data sources, incorporating exposure-led protectiveness (e.g., toxicity-exposure ratios for environmental compartments).

3. Tiered IATA Workflow

Tier 0: Exposure and Prioritization (No Testing Needed for Low-Exposure Scenarios)

- Use environmental exposure waiving (low release, rapid degradation) or Threshold of Toxicological Concern (TTC) adapted for ecotox.

- Read-across/category approach from analogs (e.g., OECD QSAR Toolbox with ecotox profilers).

Outcome: Low concern → no further testing; proceed otherwise.

Tier 1: In Silico Predictions and Environmental Fate

- Physicochemical and fate properties (e.g., OPERA, EPI Suite for logKow, biodegradation).

- Bioaccumulation modeling (e.g., BCFBAF, Arnot-Gobas models).

- (Q)SAR for ecotox endpoints (e.g., ECOSAR, VEGA, DEMETRA for algae/daphnia/fish LC50/EC50).

- Persistence/degradation predictions (e.g., BIOWIN, CATALOGIC %BOD).

- ***Outcome:*** Predict baseline toxicity (narcosis), specific MoA, and P/B potential; flag high-concern structures (e.g., high logKow + low degradability).

Tier 2: High-Throughput In Vitro/Alternative Organism Bioactivity Screening

Broad coverage assays: Baseline toxicity models (e.g., RTgill-W1 fish gill cell line cytotoxicity).

Targeted ecotox-specific assays:

- ***Algae:*** Green algae growth inhibition surrogates (e.g., photosynthetic efficiency in vitro).

- ***Invertebrates:*** Daphnia immobilization/reproduction models (e.g., embryonic cell lines, eRNAi if applicable).

- ***Fish***: Fish embryo acute toxicity test (FET, OECD TG 236 – non-protected stages), eleutheroembryo assays, zebrafish embryo/larval endpoints (non-animal lifecycle stages).

- ***Mechanistic:*** High-content assays for key events (e.g., oxidative stress, endocrine disruption in fish cell lines).

Exposure simulation: Static/renewal in vitro systems; chronic-relevant repeated dosing.

PoD derivation: Benchmark concentration (BMC) or ECx from concentration-response curves.

Outcome: Bioactivity PoD (e.g., lowest EC50); classify as baseline/specific toxicity, low/medium/high potency.

Tier 3: Advanced Integration and Refinement

- ***QIVIVE/QIVIV Extrapolation:*** Convert in vitro/alternative PoDs to whole-organism equivalents via SSD modeling or acute-to-chronic ratios.

- ***AOP/MoA-informed WoE***: Link bioactivity to key events (e.g., via Eco-AOP networks for narcosis or specific mechanisms).

- ***Omics integration:*** Transcriptomics for MoA classification (e.g., fish cell gene signatures).

Uncertainty factors: Apply assessment factors based on coverage gaps (e.g., 10-1000 for PNEC derivation).

Outcome: Compartment-specific PNEC (e.g., µg/L for aquatic); PBT/vPvB conclusion; classification into low/medium/high environmental concern.

Tier 4: Case-Specific Targeted Testing (If Needed)

Additional NAMs only if critical gaps (e.g., advanced microbial function assays, enhanced SSD from larger analog sets, emerging 3D models, or refined fate modeling); otherwise, rely on existing monitoring data or field studies.

Feedback loop: Refine models with new NAM data or open literature.

4. Decision Rules and Outputs

Classification Example:

- ***Low concern***: PNEC >> PEC or no bioactivity → no environmental classification/waiver.

- ***Medium***: PNEC supports safe environmental concentration.

- ***High***: Strong signals (e.g., low EC50 + high BCF) → classification/risk management.

PoD for Risk Assessment: Conservative PNEC derived via SSD or assessment factors, using toxicity-exposure ratios.

5. Case Studies

Exemplar chemicals: e.g., surfactants (baseline toxicity proof-of-concept), PFAS (persistence/ bioaccumulation), neonicotinoids, or pharmaceuticals from ECHA/OECD ecotox IATA cases.

Demonstrate concordance with historical experimental data using NAM-only approaches.

6. Implementation Roadmap

- ***Short-term:*** Pilot for specific classes (e.g., neutral organics or readily biodegradable compounds).
- ***Mid-term:*** Harmonize with OECD/ECHA/EFSA/PARC (e.g., expanding fish alternative acceptance).
- ***Training:*** Regulator/industry workshops.

Conclusion

This open-source framework, could assist researchers to accelerate the shift to 100% animal-free toxicology testing. By focusing on human-relevant, cutting-edge New Approach Methodologies (NAMs), it addresses the toughest challenges in regulatory acceptance, suggesting structured, actionable tools that could enhance predictive accuracy, reduce costs, and promote ethical science.

Designed for discussion and collaboration, it can help bridge theory and practice, enabling you to submit robust, NAM-based data to regulators like the FDA, EMA, and OECD.

Below is a concise summary of its core areas, highlighting how each supports innovative research while aligning with global trends like the FDA's 2025 Roadmap.

1. Regulatory Landscape and NAM Acceptance

- **Overview**: Details the current status of NAMs (e.g., in silico models like QSAR/ML, organ-on-chips, AOPs/IATAs) across FDA, EMA, and OECD, with maturity rankings and industry adoption rates.
- **Appeal to Researchers**: Provides a suggested ready-reference table for selecting validated tools, saving time on literature reviews. Use it to justify NAM choices in grants or papers, leveraging high-adoption methods (e.g., >90% pharma use for in silico screening) for faster, more reliable hazard predictions.
- **Key Benefits**: Identifies gaps in chronic endpoints, to help you guide your experiments toward regulatory-compliant innovations.

2. NAMs vs. Traditional Animal Testing Comparison

- **Overview**: Side-by-side analysis across aspects like human relevance, ethics, speed/cost, mechanistic insight, and regulatory acceptance.
- **Appeal to Researchers**: Demonstrates NAM superiority (e.g., better human translation, detailed molecular data via omics)—ideal for building compelling arguments in publications or proposals. Highlights how NAMs overcome animal testing's ~90% failure rate in drug translation.
- **Key Benefits**: Empowers ethical research by quantifying advantages like days/weeks turnaround vs. months/years, wherever it is viable, enabling high-throughput studies without animal welfare concerns.

3. Challenging Endpoints and Replacement Barriers

- **Overview**: Identifies "hard-to-replace" areas (e.g., repeated-dose systemic toxicity, DART, carcinogenicity, neurotoxicity, immunotoxicity, inhalation toxicity, ecotoxicology) and root causes like validation hurdles.
- **Appeal to Researchers**: Pinpoints high-impact opportunities for your work—focusing on these to help you contribute to cutting-edge solutions, such as developing AI models or organoids that address systemic interactions.
- **Key Benefits**: Offers a way to target under-served endpoints, positioning your research at the forefront of NGRA (exposure-led, hypothesis-driven assessments).

4. Framework Building Blocks

- **Overview**: Core components including tiered IATA structures, NGRA principles, integration of tools like organoid batteries, AI PK/PD modeling, IVIVE, and weight-of-evidence (AOP-informed with uncertainty analysis).
- **Appeal to Researchers**: Suggests practical blueprints for designing experiments. Adapt them to test novel hypotheses, e.g., combining multi-organ chips with PBK modeling for precise PoD derivation.
- **Key Benefits**: Streamlines workflow from in silico predictions to bioactivity screening, reducing trial-and-error and fostering interdisciplinary collaboration.

5. Endpoint-Specific IATA Frameworks

- **Overview**: Detailed, tiered guides for key endpoints (e.g., Repeated-Dose Systemic Toxicity IATA, DART IATA), with exposure-based waiving, in vitro assays, and risk assessment trees—fully animal-free.
- **Appeal to Researchers**: Customize for specific chemicals or biologics, generating publishable data that advances fields like drug safety or environmental toxicology.
- **Key Benefits**: Fills gaps in existing roadmaps (e.g., FDA/OECD) with suggested hands-on implementation, enabling confident, data-driven submissions without defaulting to animals.

6. Implementation Roadmap and Collaboration

- **Overview**: Phased plan (short-term pilots for classes like monoclonal antibodies; mid-term harmonization; long-term full replacement with training).
- **Appeal to Researchers**: Invites you to "build upon" the framework—critique, adapt, and share. Supports the call for global efforts to validate NAMs, potentially co-authoring updates or case studies.
- **Key Benefits**: Aligns with ethical goals (sparing millions of animals) while boosting career impact through regulatory relevance and faster innovation.

Appendix 1:

Tier 2: High-Throughput In Vitro Bioactivity Screening

Tier 2 in the Tiered Integrated Approach to Testing and Assessment (IATA) Workflow for Developmental and Reproductive Toxicity (DART) focuses on high-throughput screening using in vitro (cell-based or non-animal) assays to evaluate a chemical's bioactivity. This tier builds on Tier 1's in silico predictions by experimentally assessing potential disruptions to biological pathways relevant to reproduction and development.

It aims to identify chemicals with low, medium, or high potency for DART effects without relying on animal testing. The assays cover broad mechanisms like endocrine disruption and targeted DART endpoints, incorporating advanced techniques for more realistic exposure simulations. The outcome is a bioactivity Point of Departure (PoD), such as the lowest AC50 (concentration causing 50% of maximum activity), which helps classify the chemical's potency.

Broad Coverage Assays

These provide an initial, wide-ranging scan of potential bioactivity across multiple pathways.

ToxCast/Tox21 Battery: This is a suite of high-throughput screening assays developed by the U.S. EPA and National Toxicology Program to evaluate chemicals for endocrine disruption, including estrogen receptor (ER), androgen receptor (AR), and thyroid receptor (TR) activities. ToxCast includes over 700 assays testing thousands of chemicals, while Tox21 focuses on ~10,000 compounds using automated robotics.

For DART, these assays measure how chemicals interact with hormone receptors or pathways like steroidogenesis, which can lead to reproductive or developmental issues. For example, computational models integrate data from 11 ToxCast/Tox21 assays to predict AR pathway activity by mapping to key events in adverse outcome pathways (AOPs).

They help identify potential endocrine disruptors by quantifying activation or inhibition, with results processed through pipelines like tcpl (ToxCast pipeline) for dose-response analysis. In DART contexts, these batteries are used to profile chemicals for developmental toxicity potential, often in embryonic stem cell models.

Appendix 2:

IVIVE in Tier 3 of the Tiered IATA Workflow for DART Assessment

IVIVE stands for “In Vitro to In Vivo Extrapolation”. It is a key quantitative process in next-generation risk assessment that bridges data from in vitro (test tube/cell-based) assays to predict relevant in vivo (whole organism) effects and exposures.

In Tier 3: Advanced Integration and Refinement, IVIVE plays a central role:

> "IVIVE (In Vitro to In Vivo Extrapolation): Convert in vitro PoDs to external doses via PBK (accounting for gestational kinetics, fetal Css).

What IVIVE Does in This Context

- Tier 2 provides in vitro Points of Departure (PoDs), such as Benchmark Concentrations (BMC) or lowest AC_{50} values from concentration-response curves in high-throughput screening assays (e.g., ToxCast, zebrafish embryo tests, stem cell models).

- These in vitro PoDs represent concentrations causing bioactivity in the assay dish (often nominal or free concentrations).

- IVIVE translates these into human-relevant external doses (e.g., mg/kg/day) that would produce equivalent internal concentrations (e.g., in maternal plasma or fetal tissues) capable of triggering the same biological effects in vivo.

This enables derivation of a human-relevant PoD, such as an equivalent NOAEL (No Observed Adverse Effect Level), for risk assessment.

How IVIVE Is Performed Here

The workflow specifies using PBK/PBTK modeling (Physiologically Based Kinetic/Toxicokinetic models):

- Tools like the ‘httk R package’ (high-throughput toxicokinetics) are extended for pregnancy scenarios.

- These models incorporate gestational kinetics: time-varying physiological changes during pregnancy (e.g., increased blood volume, placental transfer, fetal compartment, altered metabolism/clearance).

Key steps:

1. Use reverse dosimetry: Start from the in vitro PoD concentration and "back-calculate" the external (oral, dermal, etc.) dose needed to achieve a matching internal dose (e.g., steady-state concentration, Css, in maternal or fetal compartments).

2. Account for placental transfer and fetal exposure, critical for Developmental and Reproductive Toxicity (DART) endpoints.

3. Refine with pregnancy-specific parameters (e.g., changes in protein binding, enzyme activity, organ blood flows).

This makes the extrapolation more protective and relevant for vulnerable windows (e.g., organogenesis).

Why It's Important in Tier 3

- Combines with other elements like AOP-informed Weight of Evidence (linking bioactivity to adverse outcomes) and omics data.

- Applies uncertainty factors (e.g., 3–10 for extrapolations, higher for vulnerable populations like pregnant women/fetuses) to yield conservative PoDs.

- Outcome: Classifies concern level (low/medium/high) and supports decisions like Bioactivity Exposure Ratio (BER = PoD / estimated internal exposure).

IVIVE reduces reliance on animal data, enhances human relevance, and has shown promise in case studies (e.g., predicting developmental toxicity of valproic acid analogues or phthalates using stem cell assays + pregnancy-adjusted models). It's a cornerstone for animal-free, mechanism-based DART risk assessment.