Toxicology 527 (2026) 154589

Contents lists available at ScienceDirect

Toxicology

journal homepage: www.elsevier.com/locate/toxicol

Artificial intelligence for predictive mixture toxicology

Jose L. Domingo [a,*], Marilia Cristina Oliveira Souza [b], Fernando Barbosa Jr. [c]

[a] Universitat Rovira i Virgili, School of Medicine, Laboratory of Toxicology and Environmental Health, Sant Llorens 21, Reus, Catalonia, 43201, Spain
[b] School of Pharmaceutical Sciences of Ribeirão Preto, Department of Biomolecular Sciences, University of São Paulo, Av. do Café s/n, Ribeirão Preto, SP, 14040-903, Brazil
[c] School of Pharmaceutical Sciences of Ribeirão Preto, Department of Clinical, Toxicological and Bromatological Analysis, University of São Paulo, Ribeirão Preto, SP, Brazil

ARTICLE INFO

Keywords:
Artificial intelligence
Chemical mixtures
Computational toxicology
Exposomics
Machine learning
Predictive toxicology

ABSTRACT

Human populations and ecosystems are continuously exposed to complex mixtures of environmental contaminants rather than to individual chemicals in isolation. These mixtures include pesticides, metals and metalloids, persistent organic pollutants, endocrine-disrupting chemicals, per- and polyfluoroalkyl substances, pharmaceuticals, plastic-associated compounds, air pollutants, nanomaterials, and numerous poorly characterized substances. Their combined effects may be additive, synergistic, or antagonistic, and are strongly influenced by dose, component ratio, timing, exposure sequence, and biological susceptibility. Experimental evaluation of all environmentally relevant mixtures is infeasible because the number of possible combinations increases combinatorially. Artificial intelligence (AI) offers a possible way to address this limitation. Machine learning, deep learning, graph neural networks, Bayesian approaches, and natural-language processing can integrate heterogeneous data, including chemical structures, molecular descriptors, toxicokinetics, high-throughput screening results, omics profiles, adverse outcome pathways, biomonitoring data, and epidemiological findings. However, the evidence base is uneven. Relatively few studies have applied AI directly to experimentally characterized mixtures, and much of the current optimism is extrapolated from single-chemical toxicology. This review distinguishes explicitly between applications demonstrated in mixtures, proof-of-concept mixture applications, and approaches whose mixture use remains prospective. It further examines the methodological requirements for mixture prediction, including dose and ratio representation, additivity reference models, applicability domains, external validation, and mechanistic interpretability, and proposes a framework for regulatory-grade implementation. Current evidence does not support autonomous AI-driven regulation of mixtures. AI should complement, not replace, experimental and expert evaluation, supporting a transition toward more predictive, mechanism-informed assessment of real-world chemical exposures.

1. Introduction

1.1. Chemical mixtures as the prevailing form of real-world exposure

In recent decades, the global production, distribution, and use of chemicals have expanded substantially. A systematic analysis of national and regional chemical inventories identified more than 350,000 chemicals and chemical mixtures registered for production and use, approximately three times previous estimates (Wang et al., 2020). This expanding chemical landscape has increased the complexity of human environmental exposures and the difficulty of evaluating their potential health effects.

The exposome concept was introduced to encompass the totality of environmental exposures from conception throughout life (Siroux et al., 2016; Wild, 2005, 2012). Modern exposome research recognizes that environmental determinants of disease cannot generally be reduced to one chemical measured at one moment (Giroux, 2025). Instead, exposures are dynamic, correlated, multidimensional, and biologically embedded through molecular changes involving gene expression, epigenetic regulation, proteins, metabolites, immune function, the microbiome, and organ physiology (Crepet et al., 2025; Vermeulen et al., 2020). Human biomonitoring provides direct evidence of this multichemical reality, and the European Human Biomonitoring Initiative has demonstrated both the value and the difficulty of converting

* Corresponding author.
E-mail address: joseluis.domingo@urv.cat (J.L. Domingo).

https://doi.org/10.1016/j.tox.2026.154589
Received 16 August 2026; Received in revised form 9 September 2026; Accepted 17 September 2026
Available online 19 September 2026
0300-483X/© 2026 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

biomonitoring results into mixture risk assessments (Luijten et al., 2023). Personal-exposure studies reinforce the same point: single ambient proxies frequently fail to represent the mixture actually reaching the individual (Bramer et al., 2024).

Additive effects are particularly important from a public-health perspective. Each component of a mixture may be present below its individual effect threshold or regulatory limit, while the combined exposure produces an appreciable biological response. Synergism, in which the observed joint effect exceeds a defined additive expectation, is claimed more often than it is rigorously demonstrated. Antagonism can also occur when one substance reduces the absorption, bioavailability, receptor binding, or downstream effects of another. Consequently, mixture toxicity cannot be classified reliably without explicit quantitative assumptions regarding the expected joint action (Backhaus and Faust, 2012; Cedergreen, 2014; Rider, 2022).

A recent review of mixture toxicity in human health described the limitations of single-substance assessment and emphasized the need to integrate exposomics, computational toxicology, modern risk-assessment strategies, One Health approach, and improved regulatory governance (Domingo and Nadal, 2026). The present manuscript addresses a distinct question: whether artificial intelligence (AI) can make the prediction and evaluation of complex mixture toxicity scientifically and operationally feasible.

1.2. Why conventional mixture assessment remains insufficient

The main challenge in mixture toxicology is not the coexistence of multiple chemicals, but the complexity of their interactions and the limits of single-substance data for predicting combined effects (Domingo, 2026a). The number of theoretically possible combinations increases extremely rapidly. Thus, 20 chemicals generate more than one million distinct combinations, and 30 chemicals more than one billion. These figures do not include variations in component ratios, doses, exposure routes, timing, duration, sequence, metabolism, life stage, lifestyle, or genetic susceptibility. Experimental testing cannot cover this combinatorial space (Jusko et al., 2026). Although component-based approaches are more adaptable, they depend on adequate toxicity and exposure data for the individual constituents and usually require assumptions regarding additivity (Domingo and Nadal, 2026; Drakvik et al., 2020; Kienzler et al., 2016).

Contemporary mixture risk assessment therefore uses pragmatic methods, including the hazard index, relative potency factors, toxic equivalency factors, reference-point and point-of-departure indices, response addition, whole-mixture comparisons, and tiered combinations of exposure and hazard information (Beronius et al., 2020). Risk assessors also face missing or uneven information. Many established contaminants have relatively extensive hazard datasets but inadequate contemporary exposure information. Conversely, emerging chemicals may be detectable by high-resolution analytical methods while lacking validated standards, toxicokinetic data, or relevant toxicity studies (Santhappan et al., 2026). Transformation products and metabolites may be overlooked even when they contribute materially to biological effects. The resulting data matrix is sparse: only a limited number of chemicals have been tested across a limited number of endpoints, doses, species, life stages, and experimental systems (Luijten et al., 2023; Muir et al., 2023; Wambaugh and Rager, 2022). These limitations explain why regulatory systems continue to rely predominantly on chemical-by-chemical assessment. The European and international scientific communities have repeatedly identified the need for better methods to assess combined exposures, improved data sharing, integration of new approach methodologies (NAMs), and more transparent decision schemes (Bopp et al., 2019; Drakvik et al., 2020; Nikolopoulou et al., 2023).

1.3. Emergence of artificial intelligence in toxicology

AI encompasses computational systems designed to perform tasks that ordinarily require aspects of human cognition, including pattern recognition, classification, prediction, language processing, and decision support. Machine learning is a major branch of AI in which algorithms infer relationships from data rather than being programmed exclusively through fixed rules. Deep learning uses multilayer neural networks capable of learning complex representations directly from molecular structures, images, sequences, text, or other high-dimensional inputs (Domingo, 2026b; Xu et al., 2021).

Applications of AI in toxicology evolved initially from cheminformatics and quantitative structure-activity relationship (QSAR) modeling. Conventional QSAR methods relate predefined molecular descriptors to a toxicological endpoint. Artificial neural networks, support-vector machines, random forests, gradient-boosting methods, graph convolutional networks, transformers, and multitask-learning architectures are now being applied to toxicity prediction, toxicokinetics, high-throughput screening, toxicogenomics, adverse outcome pathways (AOPs), and image-based phenotyping (Domingo, 2026b; Kleinstreuer and Hartung, 2024; Lin and Chou, 2022).

AI is also influencing mechanistic toxicology. Ciallella et al. (2021) developed a knowledge-based deep neural network using public high-throughput screening data to organize mechanistic information within an AOP context, while toxicokinetic modeling represents another important application (Malarvanann et al., 2026). Machine learning can assist parameter prediction, surrogate-model development, model calibration, sensitivity analysis, and extrapolation across chemicals or populations (Chou and Lin, 2023). For mixtures, these tools may eventually help identify metabolic competition, enzyme inhibition or induction, altered clearance, and target-tissue co-exposures that are not captured by external-dose addition alone.

The growing influence of AI has led some authors to describe an impending transformation of chemical-safety assessment. Kleinstreuer and Hartung (2024) argued that toxicology is sufficiently data rich for broad AI integration, while emphasizing the need for relevant data, mechanistic understanding, and appropriate oversight.

1.4. Why chemical mixtures constitute an especially appropriate problem for AI

Mixture toxicology has several characteristics that make it suitable for AI-assisted analysis (Luechtefeld and Hartung, 2025). First, the relationships between mixture composition, dose, and biological response are frequently nonlinear. Second, chemical components may interact through multiple toxicokinetic and toxicodynamic mechanisms. Third, relevant information is distributed across heterogeneous sources, including chemical structures, in vitro assays, animal studies, biomonitoring, epidemiology, omics, environmental monitoring, and scientific text. Fourth, the space of possible combinations is much larger than can be investigated experimentally (Backhaus and Faust, 2012; Lin and Chou, 2022; Muratov et al., 2020).

Machine learning can, in principle, identify relationships between mixture composition and biological response without testing every combination, learning whether particular structural features, molecular targets, concentration ratios, or pathway combinations are associated with greater-than-additive or less-than-additive effects (Esu et al., 2024; Limbu and Dakshanamurthy, 2023; Zhang et al., 2023). A conceptual workflow from exposure characterization to experimental and regulatory evaluation is shown in Fig. 1.

AI may also improve mixture prioritization. It is neither possible nor necessary to investigate all hypothetical combinations. Priority should be given to mixtures that occur in humans or ecosystems, contain chemicals with relevant potencies, affect common biological pathways, or involve vulnerable populations. Algorithms can integrate co-occurrence data from biomonitoring with predicted hazards,

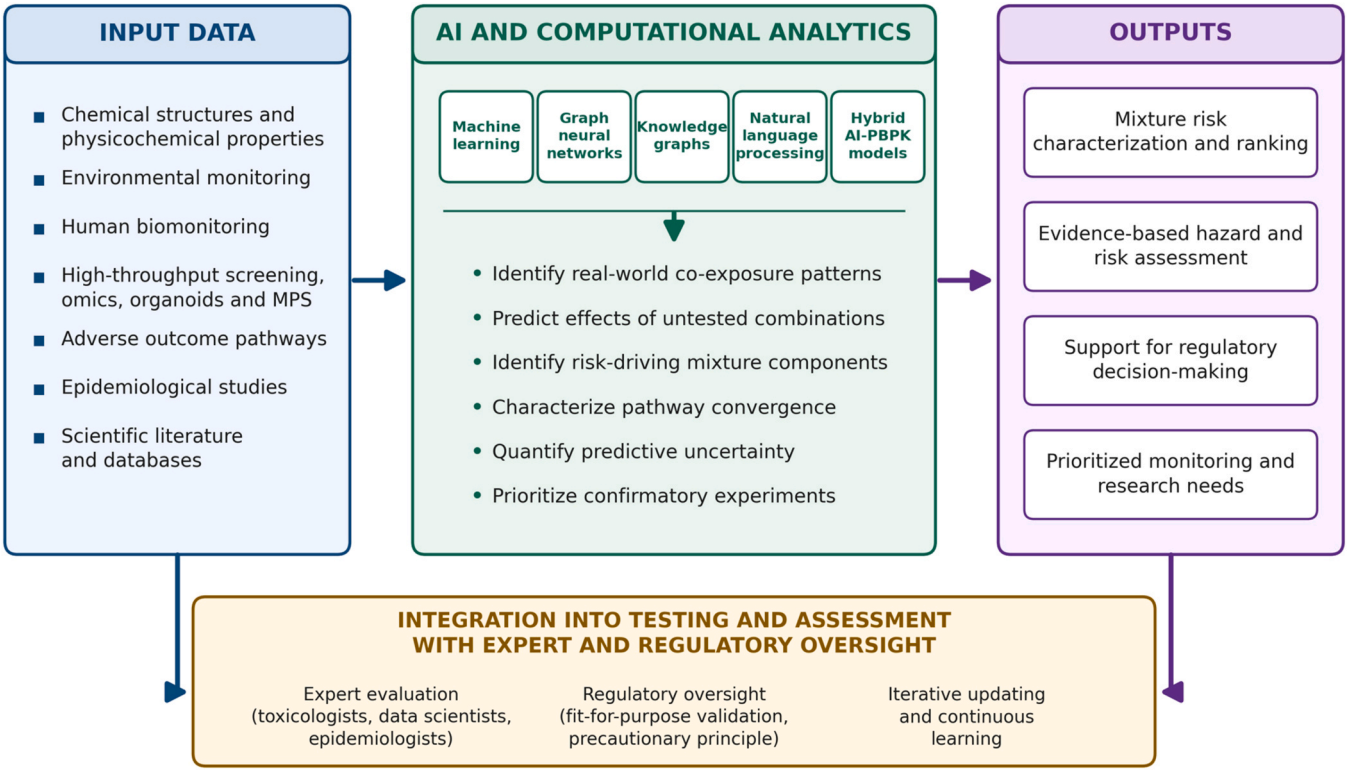


Fig. 1. Integrated artificial-intelligence framework for predictive chemical-mixture toxicology.

toxicokinetics, and mechanistic similarity to identify the mixtures most deserving experimental or regulatory attention (Braun and Escher, 2023; Reale et al., 2024; Wambaugh and Rager, 2022).

1.5. *From descriptive to predictive mixture toxicology*

Most mixture studies remain descriptive or retrospective. Investigators expose a biological system to a selected combination, observe the response, and compare it with an additive reference model. Epidemiological studies measure multiple exposures and examine their associations with health outcomes. Although these approaches remain indispensable, they evaluate only a minute fraction of potential mixtures (Backhaus and Faust, 2012; Cedergreen, 2014; Domingo and Nadal, 2026). Predictive mixture toxicology addresses a different question: given the identities, concentrations, timing, biological context, and available mechanistic information for a mixture, can its likely toxicological effects be estimated before direct testing? A useful predictive system would need to identify the principal risk-driving components, estimate whether combined effects are consistent with addition or interaction, predict relevant target organs and pathways, quantify uncertainty, and specify when the mixture falls outside the model's applicability domain.

1.6. *Objectives, scope, and novelty of the present review*

This review examines recent advances in AI-based toxicology with particular focus on assessing and predicting the harmful effects of chemical mixtures. Its intended contribution is deliberately narrower than a general survey of AI in toxicology. Previous reviews have largely addressed either conventional mixture toxicology without substantial computational content, or AI applications to single-chemical toxicity prediction. In contrast, the present review evaluates the specific methodological requirements for predicting mixture effects, namely the explicit representation of dose, component ratio, exposure duration and sequence; the choice of an additivity reference model; mixture-aware dataset partitioning and applicability domains; and mechanistic interpretability that can withstand regulatory scrutiny. It integrates these requirements across mechanistic toxicology, exposomics, NAMs, physiologically based pharmacokinetic (PBPK) modeling, and AOP frameworks, and proposes a framework for regulatory-grade implementation.

Equally important, this review is explicit about the maturity of the underlying evidence. Relatively few published studies have applied AI directly to experimentally characterized chemical mixtures, and a substantial part of the field consists of methods validated for individual substances whose extension to mixtures remains a reasonable but untested proposition. To avoid conflating technical feasibility with empirical validation, each methodological family is classified according to its current evidence status, and this classification is carried through the text, Table 1, and the conclusions.

2. Methods

2.1. *Literature-search strategy*

Given the interdisciplinary and rapidly evolving nature of the field, a narrative (non-systematic) review was conducted to provide a broad, integrative overview of the intersection of AI and chemical-mixture toxicology. A flexible, critical synthesis of concepts and emerging trends was carried out. A narrative approach was selected to integrate heterogeneous methodological, experimental, epidemiological, and regulatory evidence and to critically distinguish demonstrated applications from prospective developments. Although this heterogeneity limits the feasibility of quantitative pooling, it does not itself preclude a systematic review.

To inform this narrative synthesis, a structured literature search was performed in PubMed/MEDLINE, Scopus, Web of Science, and Google Scholar. The search focused primarily on publications from the last ten years (2016–2026), a period marked by the widespread dissemination of AI methods and included studies available through 15 July 2026. Earlier studies establishing basic concepts in mixture assessment, QSAR, PBPK

Table 1
Principal artificial-intelligence approaches, their potential applications in chemical-mixture toxicology, and the status of the supporting evidence.

AI approach	Typical input data	Potential role in mixture toxicology	Principal strength	Principal limitation	Current evidence status in mixtures
Established statistical mixture methods (WQS, BKMR, quantile g-computation)	Measured co-exposure profiles and health outcomes	Estimation of joint effects and component contributions; established comparator for AI-based mixture models.	Purpose-built for correlated multi-chemical exposures; interpretable weights	Observational; vulnerable to confounding and exposure measurement error	Demonstrated in mixtures (established comparator)
Supervised machine learning (random forest, SVM, gradient boosting)	Molecular descriptors, concentrations, component ratios, assay metadata	Prediction of mixture potency and deviation from an additive baseline	Robust with moderate dataset sizes; feature importance available	Requires curated mixture data; limited extrapolation beyond training space	Proof-of-concept in mixtures (Esu et al., 2024; Weichert et al., 2025)
Deep learning and representation learning	Structures, images, spectra, omics matrices, dose vectors	Modelling of higher-order and dose-dependent interactions	Captures complex nonlinearity; multitask and multimodal capability	Data-hungry; limited interpretability; overfitting risk	Proof-of-concept in mixtures (Limbu and Dakshanamurthy, 2023)
Machine-learning QSAR and read-across	Fingerprints, physicochemical descriptors, mixture descriptors	Two-stage prediction of component potency then deviation from additivity	Separates additive expectation from evidence of interaction	Structural similarity does not guarantee toxicological similarity	Proof-of-concept in mixtures (Zhang et al., 2023)
Knowledge graphs	Chemicals, targets, pathways, organs, outcomes	Identification of mechanistic links and convergent toxicity	Mechanistically structured and potentially interpretable	Dependent on completeness and correctness of source knowledge	Proof-of-concept in mixtures (Ren et al., 2025)
Explainable AI	Outputs of predictive models	Identification of risk-driving components and candidate mechanisms	Supports scientific scrutiny and regulatory confidence	Explanations may be correlational rather than causal	Proof-of-concept in mixtures (Chen et al., 2024)
Graph neural networks	Molecular graphs; proposed higher-level mixture graphs	Joint learning of intramolecular structure and inter-chemical relationships	Learns representations without predefined descriptors	No validated mixture implementation published to date	Extrapolated from single-chemical toxicology
Transformers and chemical foundation models	Molecular strings, text, omics, multimodal data	Knowledge transfer to data-poor mixtures; literature and data integration	Pretraining enables transfer across tasks	Opacity, training-data bias, high computational demand	Extrapolated from single-chemical toxicology
Active learning and closed-loop design	Existing experimental data and model uncertainty	Selection of the most informative mixtures for testing	Reduces experimental burden	May overlook biology not represented in the initial model	Extrapolated from single-chemical toxicology (Herman et al., 2023)
Hybrid AI-PBPK models	Chemical properties, physiology, exposure, kinetics	Prediction of internal doses and toxicokinetic interactions	Preserves mass balance and physiological constraints	Parameter uncertainty; scarce mixture kinetic data	Extrapolated from single-chemical toxicology
Multimodal AI	Structures, omics, imaging, exposure and clinical data	Integrated prediction of real-world mixture toxicity	Captures complementary evidence streams	Data harmonization and missing modalities	Extrapolated from single-chemical toxicology
Quantitative read-across structure-activity relationship (q-RASAR)	Structural/ physicochemical descriptors combined with read-across-derived similarity and error metrics	Two-stage prediction of binary/multicomponent mixture potency, combining QSAR with similarity-weighted read-across	Combines the interpretability of read-across with the predictive power of machine learning; comparatively data-efficient	Performance depends on the quality and coverage of the analogue (training) set; limited experience beyond two or three components	Proof-of-concept in mixtures (Chatterjee et al., 2023)

Note: The "current evidence status" column summarizes evidence of application to chemical mixtures according to the categories defined in Section 2.3. Application to mixtures does not, by itself, establish predictive validity for previously untested mixtures. Generalization must be evaluated separately for new combinations of known components, mixtures containing components absent from the training data, and changes in concentrations, component ratios, exposure duration, or other exposure conditions. Accordingly, neither the "demonstrated in mixtures" nor the "proof-of-concept in mixtures" designation alone establishes predictive validity beyond the conditions evaluated in the supporting studies. AOP, adverse outcome pathway; BKMR, Bayesian kernel machine regression; PBPK, physiologically based pharmacokinetic; QSAR, quantitative structure-activity relationship; SVM, support-vector machine; WQS, weighted quantile sum.

modeling, AOPs, or regulatory evaluation were also considered. The principal search equation combined three blocks with the Boolean operator AND: (i) ("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "graph neural network" OR transformer OR "large language model"); (ii) (mixture OR "combined exposure" OR "cumulative risk" OR "multiple chemical*" OR co-exposure OR multipollutant); and (iii) (toxicology OR toxicity OR "risk assessment" OR exposome OR biomonitoring OR QSAR OR PBPK OR "adverse outcome pathway" OR "new approach methodologies"). Equivalent syntax was used in each database. Reference lists of relevant articles were also screened to capture additional seminal or grey literature.

This is a narrative review, which does not report formal PRISMA-style deduplication counts or a complete study-selection flow diagram. As stated above, it is not a systematic review and does not present a quantitatively enumerated or exhaustive inventory of included studies. Its objective is a critical, illustrative synthesis of methodological principles and evidence maturity, not a bibliometric census. The principal Boolean search equation and the literature coverage cutoff are reported above. Database-specific search strings and individual search execution dates are not reported, which limits reproducibility. We consider that full PRISMA-style reporting (including documented deduplication procedures, a formal study-selection flow diagram, and inter-reviewer agreement statistics) would be most appropriately undertaken in a dedicated systematic or scoping review addressing this specific question. Such a review, which we identify as an important priority for future research, would allow these procedures to be implemented prospectively and reported transparently. Retrospectively applying these elements to the present narrative synthesis could create a misleading impression of systematic comprehensiveness and methodological exhaustiveness that was neither intended nor provided.

2.2. Eligibility criteria

Priority was given to peer-reviewed studies that developed or evaluated AI-based toxicological predictions; addressed mixtures, combined exposures, or cumulative toxicity; integrated structural, mechanistic,

toxicokinetic, exposure, omics, or epidemiological information; or examined validation, explainability, reproducibility, applicability domains, and regulatory acceptance. Studies restricted to therapeutic efficacy, drug discovery without a substantive toxicity component, or AI applications unrelated to chemical hazard or exposure were excluded.

2.3. Classification of evidence status

Because the field is uneven in maturity, each methodological family and application discussed below was assigned to one of three categories, which are also reported in Table 1. (i) Demonstrated in mixtures: the approach has been applied and evaluated on experimentally or epidemiologically characterized mixtures in more than one independent study and is already used in practice for combined-exposure analysis. (ii) Proof-of-concept in mixtures: at least one published study has applied the approach to defined mixtures, but the evidence rests on a small number of reports, typically without external or prospective validation. (iii) Extrapolated from single-chemical toxicology: the approach has been validated for individual substances, and its application to mixtures is methodologically plausible but has not yet been demonstrated. Throughout the text, statements concerning category (iii) are expressed in conditional terms, and are presented as research proposals rather than as available capabilities.

To reduce subjectivity in applying these three categories, classification was based on explicit criteria rather than qualitative impression. An application was assigned to "demonstrated in mixtures" only if it met all of the following: application to a mixture of two or more defined chemical components with specified concentrations or ratios, a stated comparator or reference model (e.g., concentration addition or independent action); replication across more than one independent published study, and evidence of external or prospective validation, or established regulatory or applied use. "Proof-of-concept in mixtures" required application to a defined mixture with reported concentrations or ratios but permitted evidence from a single published study without external validation. "Extrapolated from single-chemical toxicology" was assigned when a method had documented single-chemical validation but no identified mixture-specific application meeting the above criteria at the time of the literature search (July 15. 2026). These explicit criteria were applied to the classifications summarized in Table 1.

2.4. Evidence evaluation and synthesis

Selected studies were evaluated for methodological relevance, data quality, model validation, biological interpretability, and relevance to chemical mixtures. It was examined whether they used separate training, validation, and independent test datasets; avoided chemical or experimental overlap between datasets; reported doses, concentrations, and component ratios; defined model applicability domains; compared AI results with conventional models; quantified uncertainty; included mechanistic or explainability analyses; conducted external or prospective validation; and provided sufficient access to data or code to support reproducibility. Because mixture-specific AI studies remain limited, evidence from single-chemical predictive toxicology was included only when the methods were directly transferable, and such applications are identified as prospective rather than established regulatory practice.

2.5. Defining the mixture-prediction problem

Because "AI for mixture toxicity" can refer to several distinct scientific tasks, this review distinguishes explicitly among six prediction targets that recur throughout the following sections: (1) prediction of a toxicological endpoint for a fully defined mixture (composition, concentrations, and ratios known), (2) prediction of the magnitude and direction of departure from an additive reference model, (3) identification of the risk-driving components within a complex mixture, (4) prioritization of untested mixtures or components for experimental testing, (5) quantitative risk characterization suitable for regulatory decision-making, and (6) causal epidemiological inference regarding health effects of real-world co-exposures. These tasks differ substantially in required inputs, appropriate outputs, validation standards, causal assumptions, and regulatory consequences. A method that performs well for one task (for example, prioritization) should not be assumed fit for a more demanding task (for instance, quantitative risk characterization) without task-specific validation. This distinction is applied throughout the review and is reflected in the evidence-status classification above and in the fit-for-purpose requirements summarized in Table 2.

3. Artificial-intelligence methodologies applicable to chemical-mixture toxicology

AI is not a single methodology but a broad family of computational approaches differing in their assumptions, data requirements, interpretability, and predictive capabilities. Selecting an appropriate algorithm for mixture toxicology depends on the scientific question under investigation. Unsupervised learning can reveal recurrent co-exposure patterns without predefined outcome labels, whereas graph-based and multimodal models may integrate chemical structures with molecular targets, omics responses, exposure information, and phenotypic data (Alam et al., 2025; Jusko et al., 2026). The major methodological families, together with their mixture-specific strengths, limitations, and current evidence status, are summarized in Table 1. No algorithm can compensate for an inappropriate study design or poor-quality input data. For mixture applications, model development must account explicitly for chemical identities, concentrations, component ratios, exposure duration, biological model, endpoint definition, and the reference model used to classify additivity or interaction (Gant et al., 2026; Kleinstreuer and Hartung, 2024; Lin and Chou, 2022).

3.1. Conventional supervised machine learning

Supervised machine-learning algorithms learn a mathematical relationship between input features and known outcomes. Commonly used methods include random forests, support-vector machines, gradient-boosting algorithms, elastic-net regression, k-nearest-neighbor models, and artificial neural networks. These approaches have been applied extensively to single-chemical toxicity prediction, including genotoxicity, hepatotoxicity, cardiotoxicity, developmental toxicity, skin sensitization, endocrine activity, mitochondrial toxicity, and acute systemic toxicity (Banerjee et al., 2024a; García de Lomana et al., 2023; Jeong and Choi, 2022; Ketkar et al., 2023; Thienpont et al., 2023).

In toxicology of mixtures, the input matrix may contain several categories of predictors. Chemical-level features include molecular fingerprints, physicochemical properties, structural alerts, predicted molecular targets, and toxicokinetic parameters. Mixture-level features include the number of components, molar fractions, concentration ratios, total administered concentration, shared targets, pathway overlap, and predicted metabolic interactions. Biological features may include species, cell type, tissue, sex, age, assay platform, exposure duration, and transcriptomic or metabolomic responses (Esu et al., 2024; Zhang et al., 2023). Supervised learning has already been applied to defined mixtures: Esu et al. (2024) derived machine-learning dose-response relationships that account for interactions among components of particulate-matter mixtures, and Weichert et al. (2025) combined in silico ecotoxicity predictions with mixture-risk modeling to identify sites at risk and the principal risk drivers along a European river. Therefore, these applications support the classification of supervised machine learning as proof-of-concept in mixtures, including a demonstrated role in ecological screening. Nonetheless, model complexity should be proportionate to the available data and decision context rather than selected because an algorithm is technologically novel (Carrico et al., 2015; Yu et al., 2022; Zhu et al., 2024).

Table 2
Minimum requirements for regulatory-grade AI models in chemical-mixture toxicology.

Domain	Minimum requirement	Preferred evidence	Regulatory concern if absent
Decision context	Clearly defined intended use	Prespecified regulatory question and performance threshold	Model may be applied beyond its validated purpose
Data provenance	Traceable source for all training and validation data	Curated data with persistent identifiers and quality ratings	Undetected errors, duplication, or unverifiable evidence
Mixture specification	Complete component identities, concentrations, ratios, duration, and sequence	Measured internal doses and verified formulations	Prediction cannot be reproduced or interpreted
Endpoint definition	Standardized and biologically relevant outcome	Quantitative concentration-response data	Inconsistent or subjective labels
Additivity reference	Explicit null model	Comparison across more than one scientifically justified model	Claims of synergy or antagonism may be meaningless
Dataset partitioning	Independence of training and test data	Scaffold-, component-, laboratory-, and temporal-split validation	Data leakage and inflated performance
External validation	Evaluation on an independent dataset	Prospective multicentre validation	Uncertain generalizability
Applicability domain	Defined chemical, mixture, biological, and exposure domains	Quantitative distance or conformity metrics	False confidence for unfamiliar mixtures
Uncertainty	Calibrated prediction intervals or probabilities	Separation of epistemic and aleatory uncertainty	Point estimates may be misleading
Mechanistic plausibility	Toxicologically credible explanation	Concordance with targets, pathways, omics, and AOPs	Correct prediction may arise from spurious correlations
Explainability	Identification of influential components and variables	Global and local explanations evaluated by toxicologists	Limited regulatory trust and auditability
Comparator	Comparison with established assessment methods	Demonstrated improvement over concentration addition or conventional models	Added value of AI remains unknown
Reproducibility	Sufficient code, model, preprocessing, and version documentation	Public repository and executable workflow	Results cannot be independently confirmed
Robustness	Sensitivity to data perturbation and missingness assessed	Stress testing and adversarial evaluation	Predictions may be unstable
Human variability	Population relevance and subgroup performance evaluated	Diverse biomonitoring and physiological data	Vulnerable groups may be inadequately protected
Governance	Accountability, version control, audit trail, and human oversight	Independent review and post-deployment monitoring	Responsibility for errors is unclear
FAIR principles	Data findable, accessible, interoperable, and reusable where possible	Standard ontologies and persistent identifiers	Poor integration and model bias
Updating	Prospective model-change protocol	Locked regulatory versions plus controlled retraining	Conclusions may change without traceability

Note: These requirements should be applied proportionately to the intended use and consequences of model error, rather than as a strict hierarchy of decision contexts. Data provenance, adequate mixture and endpoint characterization, reproducibility, appropriate uncertainty assessment, and safeguards against data leakage are relevant from exploratory research onward. For predictive models, validation should reflect the intended prediction task and applicability domain; an explicit additivity reference is required when assessing departures from additivity. As applications progress toward consequential hazard identification, quantitative risk characterization, or regulatory decision support, stronger evidence of external validity, robustness, biological relevance, and population applicability is generally needed, together with more formal documentation, independent scrutiny, version control, and monitoring. This mapping is illustrative: requirements depend on the specific task and should not be interpreted as unnecessary in earlier-stage applications. The proportionate approach is consistent with the fit-for-purpose readiness principle discussed in Section 8.1 (de Paula Souza et al., 2026). AOP, adverse outcome pathway; FAIR, findable, accessible, interoperable, reusable.

3.2. Deep learning and representation learning

Deep-learning models contain multiple computational layers capable of learning increasingly abstract representations from raw or minimally processed data (Jiang et al., 2025). Their principal advantage is the ability to model complex nonlinear relationships and extract predictive features directly from structures, images, sequences, spectra, or omics matrices. In toxicology, deep learning has been used to analyze chemical structures, high-content microscopy, transcriptomics, histopathology, and high-throughput screening data (Zhang et al., 2025).

For mixture applications, deep learning could in principle address several analytical problems that are difficult for conventional algorithms: modeling higher-order interactions involving three or more components; learning concentration-dependent relationships when dose is incorporated explicitly; integrating multiple data modalities; and predicting several endpoints simultaneously through multitask architectures that exploit biological relationships among outcomes (Chu et al., 2026; Esu et al., 2024). Direct evidence in mixtures is currently limited to proof-of-concept work. Limbu and Dakshanamurthy (2023) developed a hybrid deep-machine-learning method to predict the dose-dependent carcinogenicity of chemical mixtures, which remains one of the few published examples in which dose and mixture composition were modeled jointly.

Cell Painting represents a particularly promising high-content approach, using multiplexed fluorescent staining and automated imaging to quantify hundreds or thousands of morphological features in exposed cells. Herman et al. (2023) applied active learning to Cell Painting and transcriptomic information to expand the applicability domain of deep-learning QSAR models. That work focused on individual compounds; extending it to mixtures by selecting combinations that maximize the diversity and informativeness of observed phenotypes is a plausible, but not yet realized, next step (Garcia de Lomana et al., 2023; Herman et al., 2023).

3.3. Molecular fingerprints, QSAR, and read-across

QSAR models remain foundational to predictive toxicology (Gadaleta et al., 2024). Traditional QSAR uses descriptors such as molecular weight, lipophilicity, electronic properties, hydrogen-bonding capacity, and structural fragments to predict biological activity. Modern machine-learning QSAR can process thousands of descriptors or molecular fingerprints and identify nonlinear relationships across large chemical spaces (Gadaleta et al., 2024; Mao et al., 2021).

For mixture prediction, component-specific QSAR outputs can be combined with concentration and interaction descriptors. One strategy is to predict the potency of each constituent and then estimate an additive baseline. A second model can subsequently estimate deviation from that baseline based on structural similarity, shared targets, metabolic pathways, or experimental mixture information. This two-stage approach separates the expected contribution of individual components from evidence of interaction (Muratov et al., 2020; Zhang et al., 2023). Zhang et al. (2023) used machine-learning-driven QSAR methods to predict nanoparticle-mixture toxicity, demonstrating that descriptors representing individual materials and mixture characteristics can be integrated within a single predictive framework, and providing one of the clearest mixture-specific proofs of concept currently available.

However, similarity does not ensure equivalent toxicity. Small structural differences may substantially alter metabolism, and structurally dissimilar chemicals may converge on the same adverse pathway. Therefore, mixture read-across should combine structural, mechanistic, toxicokinetic, and exposure information, and explicit uncertainty estimates are essential when extrapolating beyond combinations represented in the training data (Gant et al., 2026; Muratov et al., 2020).

A closely related and increasingly mixture-relevant methodology is the quantitative read-across structure-activity relationship (q-RASAR) approach, which combines conventional QSAR descriptors with similarity- and error-based parameters derived from read-across predictions within a single statistical model (Banerjee et al., 2024a). Because q-RASAR explicitly quantifies chemical similarity to training-set analogues, it provides an interpretable and comparatively data-efficient alternative to purely data-driven machine learning when mixture-specific training data are scarce. Direct mixture applications are emerging. Thus, Chatterjee et al. (2023) developed a machine-learning-based q-RASAR model to predict the acute contact toxicity of binary organic pesticide mixtures in honeybees, combining structural and physicochemical descriptors with read-across-derived similarity metrics and mixing rules to predict the toxicity of untested combinations. The study illustrates how q-RASAR can extend beyond single-chemical prediction to genuinely combinatorial mixture problems and supports its classification as proof-of-concept in mixtures, complementary to the QSAR and read-across applications described above and summarized in Table 1.

3.4. Graph neural networks

Molecules are naturally represented as graphs in which atoms are nodes and chemical bonds are edges. Graph neural networks propagate information through these molecular graphs and learn representations relevant to a target property (Chen et al., 2026). This reduces reliance on predefined molecular descriptors and can identify structural features associated with a modeled property. Comparative analyses of graph-based and descriptor-based approaches for molecular property prediction have been reported (Wen et al., 2025). In toxicology, Ketkar et al. (2023) benchmarked graph-based architectures for acute-toxicity estimation, Soulios et al. (2023) developed an enhanced toxicity-prediction framework incorporating graph neural networks, and further models have extended graph-based prediction toward interpretable toxicity assessment (Chotaliya and Agrawal, 2025). All these concern individual molecules.

A chemical mixture could be represented as a graph of graphs, in which each constituent molecule is represented by its own molecular graph while a higher-level graph connects the constituents according to relative concentration, chemical similarity, predicted target overlap, metabolic relationships, or experimental evidence of interaction. Such a hierarchical representation would allow a model to learn both intramolecular structural features and inter-chemical relationships. It should be emphasized that this remains a proposal extrapolated from single-chemical applications; to the author's knowledge, no validated implementation for toxicological mixtures has yet been published. Interpretability is a further constraint. Wu et al. (2023) developed a substructure-masking explanation method that attributes graph-neural-network predictions to chemically meaningful fragments, evaluated for properties including genotoxicity and cardiotoxicity.

3.5. Transformers and chemical foundation models

Transformer architectures were initially developed for language processing but have subsequently been adapted to molecular strings, protein sequences, chemical reactions, omics data, images, and multimodal biomedical information (Madan et al., 2024). They use attention mechanisms to learn relationships among elements of an input sequence or set, including relationships separated by long distances. Chemical transformers can be pretrained on millions of molecular structures using self-supervised objectives, and the resulting representations may then be fine-tuned for toxicity prediction using smaller labeled datasets (Seal et al., 2025).

In principle, mixture composition can be represented as a set or sequence of molecular structures accompanied by concentrations, enabling attention mechanisms to learn which constituents and pairwise relationships are most relevant to the predicted endpoint (Chu et al., 2026; Yang et al., 2019). Multimodal-learning systems may be especially valuable in this context. Chu et al. (2026) developed a mixture-of-experts strategy for multimodal toxicity prediction, in which specialized sub-networks are each trained on a different data modality and a learned gating function weights their outputs, rather than forcing a single architecture to process structurally dissimilar inputs identically. Despite the terminology, that approach concerns architectural design rather than chemical mixtures, and its transfer to mixture prediction remains prospective.

3.6. Explainable artificial intelligence

Regulatory toxicology requires more than predictive accuracy. A model must provide scientifically credible reasons for its outputs, characterize uncertainty, and define its applicability domain. Explainable AI encompasses methods intended to reveal how inputs influence predictions (Achar et al., 2024; Fridkin and Bendersky, 2026). For mixtures, explainable AI must also avoid misleading simplification, because a feature-attribution score does not demonstrate a causal mechanism. If two chemicals are strongly correlated in the training data, a model may assign importance to one while the true causal agent is the other. Explanations should therefore be interpreted as hypotheses requiring toxicological evaluation rather than as definitive mechanistic proof (Chen et al., 2024; Esu et al., 2024; Gant et al., 2026). Explainable approaches have nevertheless already been applied to real mixtures: Chen et al. (2024) used interpretable machine learning to examine two groups of metals with potentially opposing renal effects in relation to renal function in middle-aged and older adults. This limitation is particularly relevant to deep-learning and graph-based models, because their internal representations often cannot be directly linked to well-defined molecular mechanisms. Feature-attribution methods can indicate the relative statistical contribution of specific variables to a prediction. Nonetheless, they do not establish a validated mode of action. Accordingly, even when an explanation appears biologically plausible, it should not be interpreted as mechanistic proof. This

interpretability limitation (rather than necessarily inadequate predictive performance) is a key reason why outputs from AI-derived models should undergo toxicologist-led mechanistic evaluation before being considered in regulatory decision-making.

4. AI applications in the assessment of mixture toxicity

4.1. Prediction of additive, synergistic, and antagonistic effects

One of the most attractive potential applications of AI is prediction of whether a chemical combination will behave according to concentration addition, independent action, synergism, or antagonism (Perez Santin et al., 2021). However, these categories are meaningful only relative to a specified null model. A mixture can appear synergistic under one reference model and additive under another, particularly when individual concentration-response curves differ in slope or maximum effect (Backhaus and Faust, 2012; Cedergreen, 2014; Esu et al., 2024). AI models should therefore be trained using quantitative measures of deviation from an explicitly defined additive expectation rather than inconsistent verbal labels extracted from publications. Suitable outcomes include the difference between observed and predicted response, combination-index values, model-deviation ratios, or posterior probabilities that the mixture differs from the selected null model (Cedergreen, 2014; Esu et al., 2024).

Therefore, a particularly useful role for AI may be to model departures from toxicological expectation rather than toxicity alone. For a given mixture, an expected response can first be defined using an established reference model, such as concentration addition, dose addition, or independent action. Deviations between observed and reference-model-predicted mixture effects can then be quantitatively characterized (Schmitt et al., 2026), providing a potentially informative target for AI-based modeling. AI approaches could be applied to identify systematic patterns in these deviations and determine which features of mixture composition, concentration, chemical structure, biological context, or exposure conditions are associated with them. In this framework, AI would not replace established toxicological principles but would instead learn the residual complexity that conventional models fail to explain. Such an approach could improve mechanistic interpretability by generating hypotheses about the conditions under which synergistic or antagonistic interactions emerge, while providing a more transparent basis for evaluating whether AI offers meaningful information beyond established mixture models.

Binary classification of synergy versus no synergy is insufficient for risk assessment. A small but statistically significant deviation may be toxicologically negligible, whereas a moderate deviation at an environmentally relevant concentration may be important (Elasan, 2024). Models should predict both the direction and the magnitude of interaction and should preserve concentration-response information (Cedergreen, 2014; Limbu and Dakshanamurthy, 2023).

Mechanistically informed features may improve prediction. These include shared targets, complementary pathway perturbations, inhibition or induction of metabolic enzymes, transporter competition, receptor cross-talk, oxidative-stress potential, mitochondrial effects, and target-organ concordance. Toxicokinetic interactions are particularly important because one chemical may alter the internal concentration of another without sharing its primary molecular target (Chou and Lin, 2023; Ciallella et al., 2021).

4.2. Identifying risk-driving components

Complex environmental samples may contain dozens or hundreds of measured chemicals, but overall biological activity can be driven disproportionately by a smaller subset. AI can support component prioritization by combining measured concentrations with predicted potency, target relevance, and model-attribution scores (Luijten et al., 2023). Experimental confirmation could then proceed through mixture-deletion or component-addition studies. In a deletion design, individual components are removed sequentially from the complete mixture to determine whether predicted drivers materially alter the biological response (Chen et al., 2024; Esu et al., 2024). This strategy could substantially reduce the experimental burden: instead of testing every possible combination, investigators could focus on the complete mixture, the predicted high-contribution subset, and a limited number of targeted deletions (Carlin and Rider, 2024). Correlated components nevertheless remain problematic, because a chemical may appear important simply because it serves as a surrogate for an unmeasured or co-occurring contaminant (Hou et al., 2025).

4.3. Prioritizing mixtures for experimental testing

AI-based prioritization should integrate three domains: evidence of co-exposure, intrinsic or predicted hazard, and biological plausibility of joint action. Mixtures consistently observed in human biomonitoring, food, drinking water, occupational environments, or ecological matrices should receive higher priority than hypothetical combinations unlikely to occur (Luijten et al., 2023; Wambaugh and Rager, 2022). Network analysis can represent co-exposure relationships, while supervised models can connect the resulting clusters with biological outcomes. Priority scores can then incorporate concentration, detection frequency, toxicological potency, vulnerable-population relevance, persistence, and predicted interaction (Bobb et al., 2015; Joubert et al., 2022; Zhu et al., 2024). Computational approaches specifically developed to predict mixture toxicity have been reviewed previously (Kar and Leszczynski, 2019).

Active learning could make testing more efficient: rather than selecting mixtures randomly or exhaustively, the algorithm identifies the untested combination expected to reduce predictive uncertainty most strongly. Herman et al. (2023) demonstrated the utility of active learning in a toxicity-prediction context integrating phenotypic information for individual compounds. For mixture toxicology, the same principle could guide selection of component ratios, concentrations, time points, and biological models, although this application has not yet been experimentally validated.

Although this review focuses on human health, comparable approaches are emerging in ecological mixture risk assessment, where the combinatorial problem is at least as severe because far fewer aquatic species than human endpoints have measured toxicity data for most environmental contaminants. Weichert et al. (2025) applied machine-learning-based in silico ecotoxicity predictions to fill hazard-data gaps along a small European river, combining multiple exposure and hazard scenarios across several taxonomic groups and identifying pharmaceuticals as the main risk drivers. This is among the clearest examples in which AI-derived predictions have been used operationally within a mixture risk assessment.

4.4. Dose-response and temporal prediction

The utility of AI for regulatory assessment depends critically on quantitative prediction across doses. Many published toxicity models classify compounds as toxic or non-toxic without specifying exposure level. Such models may support screening but cannot establish a point of departure or evaluate whether typical human exposure approaches a biologically active concentration (Yang et al., 2018). Neural ordinary differential equations, recurrent neural networks, temporal convolutional networks, and transformer models may be used to analyze dynamic biological responses (Chou and Lin, 2023; Limbu and Dakshanamurthy, 2023; Sun et al., 2026).

Time is particularly important for mixtures (Ashauer et al., 2017). Simultaneous exposure can produce a different effect from sequential exposure if the first chemical induces metabolic enzymes, depletes antioxidant reserves, damages a barrier, or changes receptor expression before the second exposure. A static mixture representation cannot

capture these processes. Longitudinal models should therefore distinguish component identity, concentration, sequence, and timing (Backhaus and Faust, 2012; Domingo and Nadal, 2026).

4.5. AI-assisted PBPK and toxicokinetic modeling

PBPK models translate external exposure into predicted concentrations in blood, tissues, and target organs. For mixtures, PBPK models can represent enzyme competition, transporter interactions, protein-binding displacement, altered absorption, and shared elimination pathways (Gonnabathula et al., 2025). Surrogate models can approximate computationally intensive PBPK simulations, enabling rapid evaluation of thousands of exposure scenarios. Chou and Lin (2023) reviewed the growing role of machine learning in PBPK modeling and identified parameter prediction, model optimization, and population extrapolation as major opportunities. Machine learning has likewise been used to predict specific toxicokinetic properties, such as the propensity of xenobiotics to transfer into human milk (Vijayaraghavan et al., 2024).

Hybrid models combining PBPK structure with machine learning may be preferable to fully data-driven approaches. The mechanistic PBPK component preserves mass balance and physiological constraints, while AI estimates poorly characterized parameters or nonlinear interactions. For mixtures, this framework could distinguish toxicodynamic interaction from apparent synergy caused by altered internal exposure (Chou and Lin, 2023; Lin and Chou, 2022). Published applications remain almost entirely single-chemical, so the mixture use of hybrid AI-PBPK models is currently extrapolated rather than demonstrated.

A further dimension of AI-assisted PBPK modeling concerns genetic variation in xenobiotic-metabolizing enzymes, transporters, and receptors. Internal exposure is not simply a function of external dose. It is a biological phenotype shaped by genetic variation in cytochrome P450 (CYP) enzymes and other determinants of absorption, distribution, metabolism, and elimination, so that two individuals exposed to an identical external mixture can experience materially different tissue concentrations, metabolite profiles, and downstream biological responses (Grzegorzewski et al., 2022). Mechanistic PBPK models incorporating CYP genotype or activity-score information have already been developed and validated for individual xenobiotics (Grzegorzewski et al., 2022), showing the feasibility of embedding pharmacogenomic variability within a physiologically structured framework. Although extending this approach to AI-augmented PBPK models for mixtures, in which genomic variation is treated as an explicit input alongside exposure history and mixture composition, remains prospective, it is methodologically plausible.

In a mixture context, a genetic variant affecting one metabolic or transport pathway could simultaneously alter the internal concentration of one component and the metabolic disposition of another. Thus, an interaction that appears synergistic or antagonistic at the population level may, in some individuals, primarily reflect genotype-dependent toxicokinetic interaction rather than a toxicodynamic one (Chou and Lin, 2023). Rather than treating genetic susceptibility only as an epidemiological effect modifier, AI-augmented PBPK models could, in principle, be used to examine how specific genetic variants alter internal exposure, integrating genomic, toxicokinetic, and tissue-dosimetry information within a single mechanistically constrained platform. Large language models may contribute to this effort not as replacements for validated mechanistic or statistical models, but as an interactive interface through which investigators can iteratively query heterogeneous genomic, exposomic, and toxicokinetic datasets and formulate hypotheses (for example, exploring how alternative metabolic genotypes might alter tissue-level area-under-the-curve estimates, or whether an apparent mixture interaction persists after accounting for genotype-dependent internal dose) (Bell et al., 2026). This direction would extend the mechanistic-interpretability emphasis of this review from explaining why toxicity differs across mixtures and exposure scenarios to explaining why it differs across individuals. It is consistent with the broader observation that unresolved mixture-related health risks in genetically heterogeneous populations, such as those historically associated with contaminated drinking water at Camp Lejeune, can partly reflect this unmodeled toxicokinetic heterogeneity (Pressman and Hayes, 2026). This remains, at present, a conceptual extension rather than a demonstrated application, and is identified as a specific priority for future research.

4.6. AI and adverse outcome pathways

AOPs organize mechanistic evidence from a molecular initiating event through intermediate key events to an adverse outcome. They offer a biologically interpretable framework for grouping chemicals and evaluating whether different components converge on the same toxicological process (Ciallella et al., 2021; Mortensen et al., 2025). Ciallella et al. (2021) demonstrated a knowledge-based deep-neural-network approach that used high-throughput screening data to reveal AOP relationships. For mixtures, AOP networks are more relevant than isolated linear pathways, because several chemicals may activate different molecular initiating events that converge on a shared key event, or a single chemical may affect multiple pathways simultaneously. Graph-based AI could estimate cumulative pathway perturbation and identify network bottlenecks at which combined effects become biologically amplified (Mortensen et al., 2025).

AOP-based mixture prediction remains constrained by incomplete pathway coverage and uncertain quantitative relationships between key events. Many AOPs are qualitative, whereas regulatory prediction requires dose-, time-, and context-dependent information (Ortega-Vallbona et al., 2024). AI should therefore support, rather than substitute for, quantitative AOP development.

4.7. Multicomponent formulations and impure commercial products

The mixture-toxicity framework described above is directly transferable to a related and practically important scenario: commercial products and technical-grade materials that are formulated or manufactured as multicomponent systems instead of single, chemically pure substances. Pesticide formulations, industrial solvents, personal-care products, food-contact materials, and many technical-grade industrial chemicals routinely contain intentionally added co-formulants, solvents, stabilizers, or surfactants together with unintentional synthesis by-products and process impurities, so that the toxicologically relevant unit is frequently the formulated product rather than the nominal active ingredient (Cedergreen, 2014). AI-based mixture approaches, including the two-stage QSAR/read-across and q-RASAR strategies, are in principle well suited to this problem. They can incorporate the declared composition of a formulation together with structural and physicochemical descriptors of its individual constituents to estimate whether whole-product toxicity departs from that expected from the active ingredient alone. This approach is directly analogous to component-based mixture assessment but applied to manufactured products rather than to environmental co-exposures. A substantial additional difficulty specific to commercial and technical-grade samples is that impurity profiles are frequently unknown, proprietary, or variable between batches. Unidentified impurities cannot be represented as structural descriptors. Therefore, they are invisible to component-based AI models unless untargeted chemical characterization is concurrently performed. Consequently, AI-predicted toxicity for a declared formulation should be interpreted as an incomplete estimate of potential hazard, subject to uncertainty from uncharacterized constituents, while confirmatory whole-product testing remains necessary whenever impurity content is not fully characterized (Muir et al., 2023).

5. Artificial intelligence, exposomics, and environmental-mixture epidemiology

5.1. Untargeted chemical analysis

Exposomic platforms using high-resolution mass spectrometry can detect large numbers of exogenous chemicals and endogenous metabolites, but many features remain unidentified or only tentatively annotated (Talavera Andujar and Schymanski, 2026). The analytical challenge is no longer restricted to detecting signals; it includes identifying compounds, estimating concentrations, distinguishing exposure biomarkers from biological responses, and linking complex profiles to health outcomes (Jusko et al., 2026). AI can improve spectral matching, fragmentation prediction, retention-time prediction, and prioritization of candidate structures (Sille et al., 2025; Vermeulen et al., 2020). These capabilities are directly relevant to mixture toxicology because incomplete chemical identification causes substantial exposure misclassification. Sille et al. (2025) emphasized the role of AI-based retention-time prediction in improving chemical identification and supporting future large-scale exposome initiatives. AI may also help estimate concentrations when authentic reference standards are unavailable, although predicted quantities remain more uncertain than validated targeted measurements (Muir et al., 2023; Sille et al., 2025).

Human biomonitoring can serve as an empirical anchor for defining which chemical mixtures are present within human populations and therefore which mixtures should be prioritized for AI-assisted toxicological assessment.

5.2. Integrating exposure profiles with multi-omics

Exposome studies generate thousands of chemicals, metabolomic, transcriptomic, epigenomic, proteomic, and microbiome variables that are highly dimensional, correlated, and frequently incomplete (Wan et al., 2025). AI can identify latent exposure-response patterns and potentially distinguish biomarkers of external exposure from downstream biological effects (Guimbaud et al., 2024; Vermeulen et al., 2020).

Multimodal architecture can integrate different omics layers while preserving modality-specific information. For example, chemical exposure profiles can be related with transcriptomic pathway activation and metabolomic evidence of oxidative stress or mitochondrial dysfunction (Rogers et al., 2025), potentially revealing mechanistic signatures of mixture exposure that are not attributable to a single measured contaminant. Ren et al. (2025) illustrated this approach by coupling a biological knowledge graph with machine learning to link a mixture of per- and polyfluoroalkyl substances (PFAS) and metals to early pregnancy loss, demonstrating how mechanistic knowledge can be embedded directly into a mixture-epidemiology model rather than treated only as post hoc interpretation. This is a genuine mixture application, although it currently rests on a single study. Multi-omics models are vulnerable to overfitting, batch effects, and reverse causation, because biological profiles may reflect disease processes, medication, diet, age, or other factors rather than exposure alone. External replication and temporally resolved sampling are therefore essential (Gant et al., 2026).

5.3. Environmental epidemiology

Machine learning complements but does not replace established mixture-statistics methods. Weighted quantile sum regression, Bayesian kernel machine regression, quantile g-computation, and related approaches were developed specifically to evaluate multiple correlated exposures (Gao et al., 2025). Their strengths include explicit estimation of mixture effects and, in some cases, nonlinear or non-additive relationships (Bobb et al., 2015; Carrico et al., 2015; Keil et al., 2020; Pan et al., 2025; Yu et al., 2022; Zhu et al., 2024). These methods constitute the one area in which quantitative multi-chemical modeling is genuinely established rather than prospective, and they should be regarded as the comparator against which AI approaches are judged. AI can extend them by accommodating higher-dimensional exposure spaces, complex interactions, and unstructured data.

Predictive accuracy is not equivalent to causal inference. A model may accurately predict an outcome from a chemical mixture because exposures serve as proxies for socioeconomic conditions, occupation, diet, or geography (Chen et al., 2024; Joubert et al., 2022; Zhang et al., 2024). Causal machine learning represents a promising but demanding development, and methods such as causal forests, doubly robust learners, targeted machine learning, and Bayesian causal models may estimate heterogeneous exposure effects under specified assumptions (Komura et al., 2026). In mixture epidemiology, these approaches could examine whether combined exposures have different effects across age, sex, genetic background, nutritional status, or socioeconomic circumstances (Joubert et al., 2022; Moccia et al., 2024; Zhu et al., 2024).

The validity of causal estimates nevertheless depends on adequate confounder measurement, positivity, temporal ordering, and model specification. AI cannot recover unmeasured causal information simply from dataset size, so mechanistic toxicology, epidemiology, and exposure science must be integrated rather than treated as interchangeable evidence streams (Keil et al., 2020; Yu et al., 2022).

6. Artificial intelligence and new approach methodologies

NAMs are essential to any realistic strategy for mixture assessment. High-throughput screening, organoids, organ-on-chip systems, transcriptomic biomarkers, and high-content imaging can generate human-relevant data across multiple endpoints while reducing animal use (Collen et al., 2025; Hartung and Kleinstreuer, 2025). Collen et al. (2025) illustrated what this integration looks like in practice for a particularly data-poor endpoint, mapping strategies to combine multiple NAM batteries into a human-relevant developmental neurotoxicity testing strategy for which animal data are sparse, expensive to generate, and often only weakly predictive of the human outcome.

Machine learning can analyze these outputs, identify response patterns, estimate points of departure, and prioritize confirmatory experiments. Thienpont et al. (2023) showed how support-vector-machine and random-forest models can translate transcriptomic biomarker responses into genotoxicity predictions, while Vignaux et al. (2023) externally evaluated machine-learning predictions of acetylcholinesterase inhibition, an endpoint directly relevant to cumulative assessment of cholinesterase-inhibiting pesticides. External evaluation of this kind remains uncommon and should become standard practice.

For mixtures, NAMs can be organized as follows: (a) initial computational screening using structural alerts, target prediction, QSAR, and exposure-informed prioritization; (b) high-throughput functional testing through concentration-response assays for relevant molecular initiating events and cellular stress pathways; (c) high-content mechanistic characterization using transcriptomics, metabolomics, cell imaging, organoids, or microphysiological systems, and (d) targeted confirmation through focused in vivo, clinical, biomonitoring, or epidemiological evaluation when necessary (Drakvik et al., 2020; Hartung and Kleinstreuer, 2025; Italiya and Subramanian, 2025). AI can integrate evidence across tiers and identify discordant results. A mixture predicted to activate a pathway structurally but inactive in human-relevant assays may be deprioritized, while consistent predictions across structure, transcriptomics, and organoid responses would support further evaluation (Ciallella et al., 2021; Hartung and Kleinstreuer, 2025; Lin and Chou, 2022).

Validation remains central. Hartung and Kleinstreuer (2025) emphasized that AI-based NAMs must be assessed for reliability, relevance, reproducibility, uncertainty, and fitness for purpose. A model acceptable for preliminary screening may not be sufficient for setting a regulatory exposure limit, and validation requirements should therefore

be proportional to the consequence of the decision.

7. Interim synthesis: toward predictive mixture toxicology

AI can extend conventional mixture assessment by predicting untested combinations, identifying risk-driving components, modeling dose and time, prioritizing realistic co-exposures, integrating chemical and biological data, optimizing experiments, and quantifying uncertainty (Esu et al., 2024; Limbu and Dakshanamurthy, 2023; Singh et al., 2024; Zhang et al., 2023). The principal applications, their evidentiary requirements, and their current evidence status are summarized in Table 1.

The overall picture that emerges is deliberately unbalanced, and this asymmetry is itself an important finding. Established quantitative multichemical modeling in human populations rests almost entirely on statistical mixture methods developed within environmental epidemiology. Mixture-specific machine-learning work exists but is confined to a small number of proof-of-concept studies covering particulate-matter oxidative potential, nanoparticle mixtures, dose-dependent carcinogenicity, metal mixtures and renal function, PFAS-metal mixtures and pregnancy loss, and ecological risk along a single river. A considerable proportion of the remaining literature consists of methods validated on individual chemicals whose mixture applications are proposed rather than tested. Mixture datasets are small and heterogeneous, external validation is uncommon, and much of the current optimism is extrapolated from single-chemical toxicology (Gustavsson et al., 2023). The credible path forward is therefore a hybrid framework combining AI with exposure science, PBPK modeling, mechanistic toxicology, NAMs, conventional statistics, and expert judgment (Gant et al., 2026; Hartung and Kleinstreuer, 2025).

8. Regulatory implementation of AI-assisted mixture assessment

8.1. Current position and intended use

Regulatory authorities increasingly acknowledge that chemical-by-chemical assessment does not represent combined exposures adequately, but formal adoption of AI for mixture risk assessment remains limited. Current practice still relies principally on component-based and whole-mixture methods, concentration or dose addition, toxic-equivalency factors, hazard indices, relative potency factors, and cumulative assessment groups. AI is most defensible at present for prioritization, read-across, evidence retrieval, data-gap filling, grouping, and NAM interpretation, rather than as the sole basis for a final decision (Bell et al., 2026; Drakvik et al., 2020; Gant et al., 2026).

EFSA, the US EPA, and the OECD offer broadly compatible frameworks for assessing chemical mixtures using cumulative assessment groups, component-based methods, chemical grouping, AOPs, PBK/PBPK models, and integrated approaches to testing and assessment (IATA). AI-based tools should be evaluated for regulatory readiness and fitness for the specific decision context, including transparent validation, uncertainty characterization, and defined applicability domains (de Paula Souza et al., 2026). This distinction between general acceptability and purpose-specific readiness is now being operationalized at the European regulatory level. The Partnership for the Assessment of Risks from Chemicals (PARC) project ReadyAI is developing a readiness scoring system, built on the OECD (Q)SAR Assessment Framework, that scores AI-based models in chemical risk assessment against criteria covering data curation, model development, validation, explainability, and uncertainty quantification (de Paula Souza et al., 2026). Rather than simply accepting or rejecting a model, this scoring system indicates the decision contexts for which it is sufficiently mature. Table 2 applies the same principle to mixture assessment. A recent review has similarly emphasized transparency, validation, interpretability, and ongoing surveillance as prerequisites for regulatory acceptance of AI-based toxicology tools (Kolesnyk et al., 2026).

8.2. Fit-for-purpose validation

Validation should address technical performance, biological relevance, chemical and mixture applicability, reproducibility, uncertainty, and decision relevance. Random splits within a single dataset are inadequate when related chemicals, repeated experiments, or mixtures sharing components occur in both training and test sets. More informative evaluations include scaffold or chemical-cluster splitting, temporal and laboratory-based validation, leave-one-component-out testing, and independent prospective mixtures (Gant et al., 2026; Hartung and Kleinstreuer, 2025; Muratov et al., 2020). Mixture-specific validation should determine whether performance is retained for new combinations, unseen components, altered ratios, different concentration ranges, exposure sequences, biological systems, and population groups.

8.3. Reporting and governance

Minimum reporting should include component identities and concentrations, data provenance, preprocessing, missing-data handling, model architecture and hyperparameters, partitioning strategy, performance metrics, uncertainty methods, applicability-domain criteria, and explanation procedures. Negative results and failed external validations are essential for defining the boundaries of utility (Belfield et al., 2025; Gant et al., 2026; Muratov et al., 2020). Governance also has a societal dimension, since the acceptability of AI-supported decisions depends on public perception of the associated risks, benefits, and values (Brauner et al., 2025). The principal validation and reporting requirements are summarized in Table 2.

8.4. Implications of AI for classical mixture-assessment frameworks

The emergence of AI-based methods does not leave classical mixture-assessment frameworks unchanged. Concentration addition, independent action, the hazard index, and relative potency factors were developed under data constraints that required strong simplifying assumptions, particularly dose-additivity and toxicological similarity among components (Backhaus and Faust, 2012; Beronius et al., 2020). AI-based models challenge these frameworks in at least three ways. First, by enabling explicit, quantitative modeling of departures from additivity instead of relying on binary synergy/antagonism labels, AI can show that some interactions previously classified as synergistic can largely reflect the choice of reference model, not a genuine biological interaction. This can encourage more careful specification of the null model used in conventional assessments. Second, AI-assisted PBPK modeling may attribute apparent toxicodynamic interactions to toxicokinetic causes, such as altered internal dose, which classical component-based methods generally cannot distinguish. Third, explainable AI and knowledge-graph approaches can surface candidate mechanistic links between mixture components that were not previously considered in a classical assessment, generating hypotheses for targeted testing. At the same time, classical frameworks retain essential advantages that current AI methods cannot replicate (transparency, a long regulatory track record, minimal data requirements, and direct interpretability by non-specialist risk assessors). Consequently, the most defensible near-term relationship is not substitution but reciprocal constraint. Thus, classical additivity models provide the explicit null hypothesis against which AI predictions are evaluated, while AI helps identify the circumstances in which, and the reasons why, those classical assumptions are not met.

9. Scientific and technical limitations

Despite growing scientific interest, the translation of AI methods into routine mixture-toxicology practice has remained slow. This situation reflects identifiable technical and practical bottlenecks, rather than a

lack of scientific interest. A major challenge is the limited availability of high-quality, mixture-specific training data. Other difficulties include the explicit representation of dose, timing, and internal exposure, as well as failures of applicability when models are extrapolated beyond the distributions represented in their training data. Existing datasets may also systematically underrepresent particular chemicals, exposure scenarios, and population groups. Further concerns include the risk of interpreting predictive correlations as evidence of causality or mechanism, insufficient reproducibility and data-governance practices, and the potentially substantial computational and environmental costs of large AI models. Each of these challenges is considered below, together with its practical implications for mixture risk assessment.

9.1. Scarcity of mixture-specific training data

High-quality, quantitative data on chemical mixtures remain scarce compared with data on individual substances, and mixture studies are highly heterogeneous in composition, exposure conditions, biological systems, endpoints, and interaction definitions. Models may consequently learn study-specific features rather than generalizable toxicological principles (Drakvik et al., 2020; Gant et al., 2026; Kienzler et al., 2016). Much of the current optimism concerning AI-assisted mixture prediction is based on extrapolation from successful single-chemical models (Serra et al., 2026). Platforms such as ProTox 3.0 can predict numerous toxicity endpoints for individual chemicals, while graph neural networks, chemical-language models, and deep-learning systems increasingly provide accurate component-level predictions (Banerjee et al., 2024b; Seal et al., 2025). These resources are valuable inputs to mixture models but cannot by themselves establish how components interact.

The creation of reference mixture datasets should become an international priority. Such datasets should include both positive and negative interactions, complete concentration-response curves, standardized controls, replicated experiments, internal-dose information where feasible, and mechanistic measurements. Environmentally realistic mixtures should be represented alongside systematically designed combinations that explore chemical and biological space (Hartung and Kleinstreuer, 2025; Joubert et al., 2022).

9.2. Dose, timing, and internal exposure

Many AI toxicity models treat hazard as an intrinsic molecular property and omit exposure concentration. This approach can support prioritization but not risk characterization. For mixtures, the omission is particularly serious, because interaction may occur only within a narrow range of doses or component ratios (Limbu and Dakshanamurthy, 2023; Lin and Chou, 2022).

9.3. Applicability-domain failure and distribution shift

An AI model is generally reliable only within the chemical, biological, and experimental domain represented in its training data. Real-world mixture assessment frequently requires extrapolation to new chemicals, unusual combinations, low-dose chronic exposure, developmental windows, or susceptible populations (Hartung and Kleinstreuer, 2025; Muratov et al., 2020). Failure can also occur when an in vitro model trained on acute cytotoxicity is used to infer chronic endocrine, neurodevelopmental, immunological, or ecological effects. Applicability-domain analysis should therefore operate at several levels: chemical domain (structural and physicochemical similarity); mixture domain (similarity in components and component ratios); biological domain (assay, species, tissue, life stage, and endpoint); exposure domain (dose, duration, timing, and route); and mechanistic domain (molecular targets and biological pathways).

9.4. Bias and representativeness

Training data can reproduce longstanding biases in toxicology. Frequently studied chemicals, commercially important compounds, pharmaceuticals, and well-funded regulatory endpoints are overrepresented, whereas transformation products, low-production-volume chemicals, contaminants affecting disadvantaged communities, and chemicals used primarily in lower-income countries may be underrepresented (Gant et al., 2026; Kleinstreuer and Hartung, 2024; Luechtefeld and Hartung, 2025). Karasaki et al. (2024) evaluated machine-learning models predicting drinking-water contaminant violations across California water systems and found that model performance was systematically worse for smaller, under-resourced systems, which disproportionately serve low-income and minority communities. Performance should therefore be reported by subgroup rather than only as an aggregate metric.

9.5. Correlation versus causality

Machine learning can identify highly predictive exposure patterns without establishing that the measured chemicals cause the outcome. In epidemiological datasets, chemical mixtures may correlate with occupation, diet, socioeconomic status, residential location, medical treatment, or other unmeasured determinants (Yu et al., 2022; Zhu et al., 2024). AI cannot resolve unmeasured confounding through algorithmic complexity alone, so mechanistic evidence, prospective studies, natural experiments, toxicokinetic plausibility, and experimental confirmation remain necessary. Large language models also have difficulty weighing conflicting causal claims and may produce fluent but insufficiently critical syntheses. AI-assisted evidence retrieval and synthesis can help compare studies and identify methodological disagreements, but human evaluation remains essential, particularly when the evidence is heterogeneous or politically consequential (Bell et al., 2026; Gant et al., 2026).

9.6. Reproducibility and data governance

AI models can be difficult to reproduce when code is unavailable, datasets are proprietary, preprocessing steps are undocumented, random initialization is not controlled, software versions change, or reporting is incomplete. Generative AI systems introduce additional challenges because their outputs can change following model updates even when the prompt remains identical (Belfield et al., 2025; Gant et al., 2026; Hartung and Kleinstreuer, 2025). FAIR principles (findable, accessible, interoperable, and reusable) are particularly important for mixture toxicology (Cronin et al., 2026). Chemical identifiers, units, assay ontologies, endpoint definitions, and metadata must be standardized so that information can be integrated across databases. Analyses arising from the ECETOC workshop on AI in chemical risk assessment concluded that FAIR data are a prerequisite for reliable and unbiased AI support, although data quality alone does not replace governance and human oversight (Belfield et al., 2025; Gant et al., 2026; Mortensen et al., 2025; Zare Jeddi et al., 2023).

9.7. Energy demand and environmental footprint

Large AI models require computational infrastructure, electricity, cooling water, and hardware resources (Domingo and Shomar, 2026). Although predictive toxicology may reduce animal use and unnecessary experimentation, its environmental benefits should not be assumed automatically. Model size should be proportionate to the task, and smaller interpretable models should be preferred when they perform comparably to resource-intensive architectures (Gant et al., 2026; Kleinstreuer and Hartung, 2024).

10. Future research priorities

10.1. Data infrastructure and benchmark datasets

The most urgent requirement is an international, curated mixture-toxicity data infrastructure linking standardized chemical identifiers, exact composition and ratios, concentration-response data, exposure conditions, biological models, endpoints, additive-reference models, toxicokinetic information, omics and imaging outputs, and quality-control metadata. FAIR principles are necessary but insufficient: datasets must also be biologically interpretable, versioned, and suitable for chemical- and mixture-level external validation (Belfield et al., 2025; Gant et al., 2026; Joubert et al., 2022). Benchmark datasets should include additive, synergistic, and antagonistic examples, negative findings, complete concentration-response curves, replicate experiments, and environmentally realistic mixtures. Statistical frameworks developed for the rigorous analysis of combination experiments provide a useful model for how such benchmarks might be structured (Romero-Becerra et al., 2025).

10.2. Closed-loop experimentation and hybrid models

Active-learning systems could be scaled from single-mixture prioritization to a continuous experimental design loop. Coupling these systems with high-throughput assays, transcriptomics, Cell Painting, organoids, and microphysiological systems would allow informative regions of mixture space to be mapped without exhaustive testing. Mechanistically constrained hybrid models should combine PBPK or pathway structure with machine learning so that mass balance, physiology, and causal biological constraints are retained (Hartung and Kleinstreuer, 2025; Herman et al., 2023). Both proposals are extrapolations from single-chemical work and require experimental demonstration before they can be considered available methods.

A further priority is to extend closed-loop and hybrid AI–PBPK approaches to explicitly incorporate genomic variation in metabolizing enzymes, transporters, and receptors. This would enable direct modeling of interindividual differences in internal exposure, instead of considering external dose alone. Such approach could help move the field beyond predicting whether a mixture poses a hazard, and toward explaining why its effects vary among individuals.

10.3. Foundation models and evidence retrieval

Chemical foundation models may facilitate transfer learning from large single-chemical datasets to data-poor mixtures, but their outputs require traceable sources, calibrated uncertainty, domain-specific benchmarking, and protection against training-data bias (Li et al., 2026). Retrieval-augmented language models may accelerate literature screening and extraction of chemical identities, doses, ratios, endpoints, and interaction classifications, but every extracted claim should remain linked to the source passage and be subject to human verification (Bell et al., 2026; Gant et al., 2026; Muratov et al., 2020).

11. Proposed framework for AI-assisted mixture risk assessment

A practical framework should preserve established risk-assessment

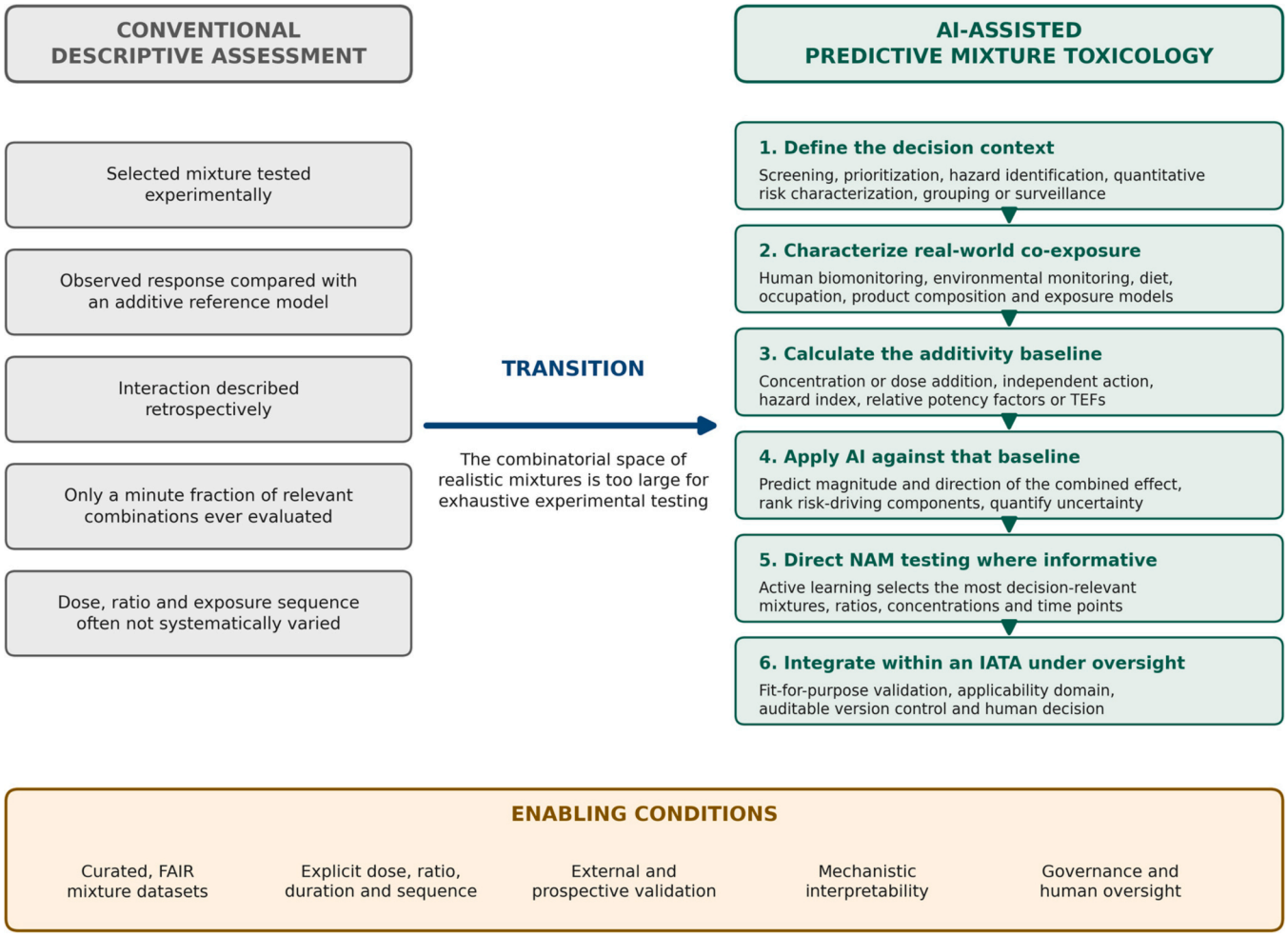


Fig. 2. Transition from conventional descriptive mixture assessment to AI-assisted predictive mixture toxicology.

principles and use AI only where it adds demonstrable value (Drakvik et al., 2020; Gant et al., 2026; Hartung and Kleinstreuer, 2025; Kienzler et al., 2016). The proposed workflow is shown in Fig. 2. First, the decision context should be defined explicitly (screening, prioritization, hazard identification, quantitative risk characterization, regulatory grouping, or surveillance), because validation requirements depend on the consequence of the decision. Second, real-world co-exposure should be characterized using biomonitoring, environmental monitoring, product-composition data, dietary and occupational information, and exposure models. Third, an established baseline, such as concentration addition, dose addition, independent action, hazard index, relative potency factors, or toxic-equivalency factors, should be calculated; AI should be evaluated against this baseline, not in isolation. Component-level evidence should then be integrated, including structures, physicochemical properties, toxicological endpoints, targets, pathways, toxicokinetics, exposure levels, and AOPs. Fourth, the model should predict the magnitude and direction of the combined response, identify risk-driving components, and quantify uncertainty, while treating dose, component ratio, duration, and timing as explicit inputs. Finally, AI-guided NAM testing should prioritize the most decision-informative mixtures and concentrations, integrating all evidence within an IATA. For high-stakes regulatory uses, models require independent review, transparent documentation and validation, prospective performance monitoring, and fully auditable version control linked to the supporting evidence. Fig. 3 shows a complementary proposed conceptual framework integrating mixture composition, exposomics, genomic and toxicokinetic variation, PBPK modeling, mechanistic evidence, and population heterogeneity to guide the development and evaluation of AI-assisted mixture risk predictions.

12. Conclusions

Chemical mixtures highlight a critical gap between real-world exposure and traditional toxicological assessment. Humans and ecosystems face ever-changing combinations of contaminants, yet most hazard evaluations and regulatory frameworks still focus on individual substances, and the number of possible combinations makes exhaustive experimental testing impractical. By integrating chemical structures, molecular targets, toxicokinetic data, high-throughput assays, omics profiles, biomonitoring measurements, environmental monitoring, and epidemiological evidence, AI may help identify the most harmful components of mixtures, predict untested combinations, uncover shared biological pathways, prioritize experiments, and assess deviations from conventional additivity assumptions. No single AI method is likely to achieve these objectives alone. Future predictive frameworks will probably require complementary integration of several of the approaches summarized in Table 1.

Nevertheless, it is essential to distinguish technical feasibility from empirical validation. Established quantitative assessment of multichemical exposures in human populations still rests mainly on statistical mixture methods from environmental epidemiology. Direct AI applications to characterized mixtures exist but remain few and largely at proof-of-concept stage, and a substantial part of the field consists of approaches validated for single substances whose mixture use is proposed rather than demonstrated. Mixture-specific datasets are scarce, interaction terminology lacks consistency, critical factors such as dose and timing are often overlooked, external validation is rare, and many models function as black boxes with limited interpretability. High accuracy in one dataset does not guarantee generalizability, mechanistic insight, or causal validity.

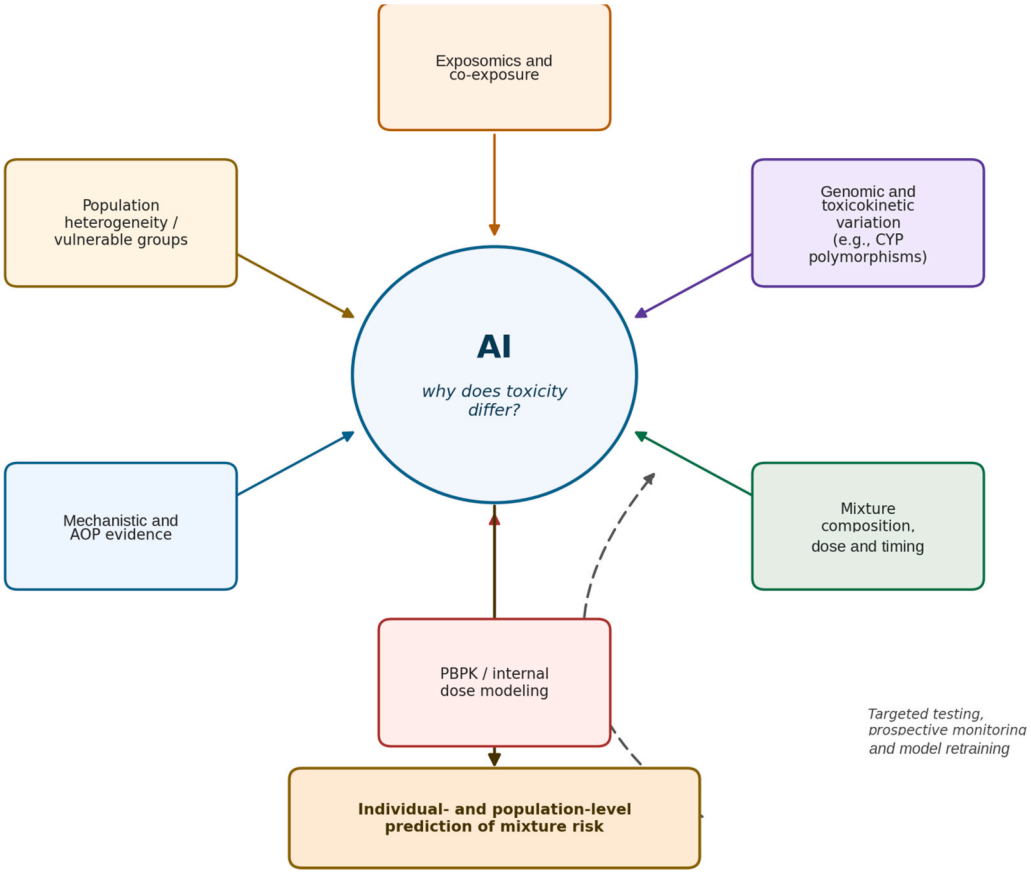


Fig. 3. Convergence of exposomics, genomic/toxicokinetic variation, mixture composition, PBPK, and mechanistic evidence within an AI-enabled framework for mixture toxicology. The dashed arc denotes the prospective-monitoring, targeted-testing, and model-retraining feedback loop linking predicted mixture risk back to model refinement.

At present, evidence does not support fully autonomous AI-driven regulation of complex mixtures. The most practical and defensible role for AI is as a decision-support tool that enhances, rather than replaces, established toxicological practice, working alongside concentration-addition models, PBPK modeling, AOPs, NAMs, human biomonitoring, epidemiology, and expert judgment. Realizing the potential of AI will depend less on algorithmic innovation than on curated and FAIR mixture datasets, explicit representation of dose, ratio and exposure sequence, mixture-aware external validation, mechanistic interpretability, and transparent governance. If those conditions are met, AI could shift the field from descriptive to predictive mixture toxicology, enabling proactive identification of hazardous co-exposures, guiding targeted experiments, supporting timely regulatory action, and improving the assessment and management of real-world chemical risks.

CRediT authorship contribution statement

Jose L. Domingo: Writing – review & editing, Writing – original draft, Supervision, Methodology, Data curation, Conceptualization. **Marilia Cristina Oliveira Souza:** Writing – review & editing, Writing – original draft, Validation, Methodology, Data curation, Conceptualization. **Fernando Barbosa:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

Figs. 1–3 were produced as schematic diagrams using conventional plotting software (Python/Matplotlib). No generative AI image-creation tool was used to create or alter any of these figures.

Funding

This research was supported in part by the Sao Paulo Research Foundation (FAPESP process numbers: 2022/06443–0, 2018/24069–3 and 2022/03476–5) and the Brazilian National Council for Scientific and Technological Development (CNPq - 406442/2022–3).

Declaration of Competing Interest

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

Data availability

This is a Review-manuscript. The information reported is based on the list of References of the paper.

References

Achar, J., Firman, J.W., Cronin, M.T.D., Öberg, G., 2024. A framework for categorizing sources of uncertainty in in silico toxicology methods: considerations for chemical toxicity predictions. Regul. Toxicol. Pharmacol. 154, 105737. https://doi.org/10.1016/j.yrtph.2024.105737.

Alam, F., Mohammed Alnazzawi, T.S., Mehmood, R., Al-Maghthawi, A., 2025. A review of the applications, benefits, and challenges of generative AI for sustainable toxicology. Curr. Res. Toxicol. 8, 100232. https://doi.org/10.1016/j.crtox.2025.100232.

Ashauer, R., O'Connor, I., Escher, B.I., 2017. Toxic mixtures in time - the sequence makes the poison. Environ. Sci. Technol. 51 (5), 3084–3092. https://doi.org/10.1021/acs.est.6b06163.

Backhaus, T., Faust, M., 2012. Predictive environmental risk assessment of chemical mixtures: a conceptual framework. Environ. Sci. Technol. 46 (5), 2564–2573. https://doi.org/10.1021/es2034125.

Banerjee, A., Kar, S., Roy, K., Patlewicz, G., Charest, N., Benfenati, E., Cronin, M.T.D., 2024a. Molecular similarity in chemical informatics and predictive toxicity modeling: from quantitative read-across (q-RA) to quantitative read-across structure-activity relationship (q-RASAR) with the application of machine learning. Crit. Rev. Toxicol. 54 (9), 659–684. https://doi.org/10.1080/10408444.2024.2386260.

Banerjee, P., Kemmler, E., Dunkel, M., Preissner, R., 2024b. ProTox 3.0: a webserver for the prediction of toxicity of chemicals. Nucleic Acids Res. 52 (W1), W513–W520. https://doi.org/10.1093/nar/gkae303.

Belfield, S.J., Basiri, H., Swapnil, C., Chrysochoou, G., Enoch, S.J., Firman, J.W., Gomatam, A., Hardy, B., Helmke, P.S., Madden, J.C., Maran, U., March-Vila, E., Nikolov, N.G., Pastor, M., Piir, G., Sild, S., Smajić, A., Spînu, N., Wedebye, E.B., Cronin, M.T.D., 2025. Moving towards making (quantitative) structure-activity relationships ((Q)SARs) for toxicity-related endpoints findable, accessible, interoperable and reusable (FAIR). ALTEX 42 (4), 657–666. https://doi.org/10.14573/altex.2411161.

Bell, S., Ackerman, M., Buck, J., Frueh, A., Maitland, J., Piotrowski, K., Sokol, M., Tussing, T.E., Zellefrow, C., McNett, M., 2026. Using artificial intelligence to enhance evidence search and synthesis. Am. J. Nurs. 126 (5), 56–60. https://doi.org/10.1097/AJN.0000000000000288.

Beronius, A., Zilliacus, J., Hanberg, A., Luijten, M., van der Voet, H., van Klaveren, J., 2020. Methodology for health risk assessment of combined exposures to multiple chemicals. Food Chem. Toxicol. 143, 111520. https://doi.org/10.1016/j.fct.2020.111520.

Bobb, J.F., Valeri, L., Claus Henn, B., Christiani, D.C., Wright, R.O., Mazumdar, M., Godleski, J.J., Coull, B.A., 2015. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostatistics 16 (3), 493–508. https://doi.org/10.1093/biostatistics/kxu058.

Bopp, S.K., Kienzler, A., Richarz, A.N., van der Linden, S.C., Paini, A., Parissis, N., Worth, A.P., 2019. Regulatory assessment and risk management of chemical mixtures: challenges and ways forward. Crit. Rev. Toxicol. 49 (2), 174–189. https://doi.org/10.1080/10408444.2019.1579169.

Bramer, L.M., Dixon, H.M., Rohlman, D., Scott, R.P., Miller, R.L., Kincl, L.D., Herbstman, J.B., Waters, K.M., Anderson, K.A.P.M., 2024. 2.5 is insufficient to explain personal PAH exposure. GeoHealth 8 (2), e2023GH000937. https://doi.org/10.1029/2023GH000937.

Braun, G., Escher, B.I., 2023. Prioritization of mixtures of neurotoxic chemicals for biomonitoring using high-throughput toxicokinetics and mixture toxicity modeling. Environ. Int. 171, 107680. https://doi.org/10.1016/j.envint.2022.107680.

Brauner, P., Glawe, F., Liehner, G.L., Vervier, L., Ziefle, M., 2025. Mapping public perception of artificial intelligence: expectations, risk-benefit tradeoffs, and value as determinants for societal acceptance. Technol. Forecast. Soc. Change 220, 124304. https://doi.org/10.1016/j.techfore.2025.124304.

Carlin, D.J., Rider, C.V., 2024. Combined exposures and mixtures research: an enduring NIEHS priority. Environ. Health Perspect. 132 (7), 75001. https://doi.org/10.1289/EHP14340.

Carrico, C., Gennings, C., Wheeler, D.C., Factor-Litvak, P., 2015. Characterization of weighted quantile sum regression for highly correlated data in a risk analysis setting. J. Agric. Biol. Environ. Stat. 20 (1), 100–120. https://doi.org/10.1007/s13253-014-0180-3.

Cedergreen, N., 2014. Quantifying synergy: a systematic review of mixture toxicity studies within environmental toxicology. PLoS. One 9 (5), e96580. https://doi.org/10.1371/journal.pone.0096580.

Chatterjee, M., Banerjee, A., Tosi, S., Carnesecchi, E., Benfenati, E., Roy, K., 2023. Machine learning - based q-RASAR modeling to predict acute contact toxicity of binary organic pesticide mixtures in honey bees. J. Hazard. Mater. 460, 132358. https://doi.org/10.1016/j.jhazmat.2023.132358.

Chen, H., Wang, M., Li, J., 2024. Exploring the association between two groups of metals with potentially opposing renal effects and renal function in middle-aged and older adults: evidence from an explainable machine learning method. Ecotoxicol. Environ. Saf. 269, 115812. https://doi.org/10.1016/j.ecoenv.2023.115812.

Chen, J., Fan, N., Lu, Y., Yang, J., Song, W., Sheng, H., Yang, Y., Chen, S., Wang, J., 2026. Intelligence on the graph: graph neural networks for mechanistic drug target discovery. J. Pharm. Anal. 16 (7), 101508. https://doi.org/10.1016/j.jpha.2025.101508.

Chotaliya, M., Agrawal, S.S., 2025. ToxiGraphNet: a graph neural network framework for precise toxicity prediction of drug molecules. J. Comput. Aided Mol. Des. 39 (2), 101. https://doi.org/10.1007/s10822-025-00683-4.

Chou, W.C., Lin, Z., 2023. Machine learning and artificial intelligence in physiologically based pharmacokinetic modeling. Toxicol. Sci. 191 (1), 1–14. https://doi.org/10.1093/toxsci/kfac101.

Chu, J.W., Park, J.H., Cho, Y.R., 2026. MEMOL: mixture of experts for multimodal learning through multi-head attention to predict drug toxicity. Comput. Methods Prog. Biomed. 273, 109088. https://doi.org/10.1016/j.cmpb.2025.109088.

Ciallella, H.L., Russo, D.P., Aleksunes, L.M., Grimm, F.A., Zhu, H., 2021. Revealing adverse outcome pathways from public high-throughput screening data to evaluate new toxicants by a knowledge-based deep neural network approach. Environ. Sci. Technol. 55 (15), 10875–10887. https://doi.org/10.1021/acs.est.1c02656.

Cöllen, E., Bartmann, K., Blum, J., Carstens, K., Celardo, I., Chatterjee, N., Corvaro, M., Dreser, N., Fritsche, E., Hartung, T., Hogberg, H.T., Knudsen, T., Koch, K., Kreutz, A., Lislien, M., Magel, V., Marty, M.S., Pallocca, G., Bal-Price, A., Rovida, C., Sachana, M., Shafer, T.J., Smirnova, L., Suciu, I., Tanaskov, Y., Tangianu, S., Wolfbeisz, C., Leist, M., 2025. Mapping out strategies to further develop human-relevant, new approach methodology (NAM)-based developmental neurotoxicity (DNT) testing. ALTEX 42 (2), 308–322. https://doi.org/10.14573/altex.2501091.

Cronin, M.T.D., Basiri, H., Belfield, S.J., Chavan, S., Chrysochoou, G., Enoch, S.J., Firman, J.W., Gomatam, A., Hardy, B., Helmke, P.S., Madden, J.C., Maran, U., March-Vila, E., Nikolov, N.G., Pastor, M., Piir, G., Popelier, P.L.A., Sild, S., Smajić, A., Spînu, N., Wedebye, E.B., 2026. The findable, accessible, interoperable, reusable (FAIR) lite principles to ensure utility of computational toxicology models. ALTEX 43 (2), 215–227. https://doi.org/10.14573/altex.2502021.

de Paula Souza, J., Blum, J., Maran, U., Sild, S., Dawson, L., Čavoški, A., Holden, L., Lee, R., Karnel, V., Meusburger, L., Fraize-Frontier, S., Walsh, A., Rivière, G., Raitano, G., Roncaglioni, A., Di Consiglio, E., Tcheremenskaia, O., Bossa, C., Wendt-Rasch, L., Puzyn, T., Fritsche, E., 2026. Advancing the implementation of artificial intelligence in regulatory frameworks for chemical safety assessment by defining robust readiness criteria. Front. Artif. Intell. 8, 1738770. https://doi.org/10.3389/frai.2025.1738770.

Domingo, J.L., 2026a. Mixture toxicity revisited: a translational review of experimental evidence from animal models to human health risk assessment. Toxicology 520, 154372. https://doi.org/10.1016/j.tox.2025.154372.

Domingo, J.L., 2026b. Artificial intelligence revolution in toxicology: clinical precision, global equity, and the 2030 roadmap. Toxicol. Lett. 418, 111871. https://doi.org/10.1016/j.toxlet.2026.111871.

Domingo, J.L., Nadal, M., 2026. Mixture toxicity in human health: integrating One Health, exposomics, and modern risk assessment strategies. Chem. Res. Toxicol. 39, 738–751. https://doi.org/10.1021/acs.chemrestox.5c00375.

Domingo, J.L., Shomar, B., 2026. Artificial intelligence's hidden environmental and public-health risks: should AI be considered a public health threat? Hum. Ecol. Risk Assess. 32 (5), 889–904. https://doi.org/10.1080/10807039.2026.2670401.

Drakvik, E., Altenburger, R., Aoki, Y., Backhaus, T., Bahadori, T., Barouki, R., Brack, W., Cronin, M.T.D., Demeneix, B., Hougaard Bennekou, S., van Klaveren, J., Kneuer, C., Kolossa-Gehring, M., Lebret, E., Posthuma, L., Reiber, L., Rider, C., Rüegg, J., Testa, G., van der Burg, B., van der Voet, H., Warhurst, A.M., van de Water, B., Yamazaki, K., Öberg, M., Bergman, Å., 2020. Statement on advancing the assessment of chemical mixtures and their risks for human health and the environment. Environ. Int. 134, 105267. https://doi.org/10.1016/j.envint.2019.105267.

Elaşan, S., 2024. The difference between clinical significance and statistical significance: an important distinction for clinical research. Turk. J. Med. Sci. 54 (6), 1419. https://doi.org/10.55730/1300-0144.5925.

Esu, C.O., Pyo, J., Cho, K., 2024. Machine learning-derived dose-response relationships considering interactions in mixtures: applications to the oxidative potential of particulate matter. J. Hazard. Mater. 475, 134864. https://doi.org/10.1016/j.jhazmat.2024.134864.

Fridkin, S., Bendersky, M., 2026. Interpretable machine learning: a comprehensive review of foundations, methods, and the path forward. WIREs Data Min. Knowl. Discov. 16 (1), e70075. https://doi.org/10.1002/widm.70075.

Gadaleta, D., García, de Lomana, M., Serrano-Candelas, E., Ortega-Vallbona, R., Gozalbes, R., Roncaglioni, A., Benfenati, E., 2024. Quantitative structure-activity relationships of chemical bioactivity toward proteins associated with molecular initiating events of organ-specific toxicity. J. Chemin.-. 16 (1), 122. https://doi.org/10.1186/s13321-024-00917-x.

Gant, T.W., Boxall, A., Burgwinkel, D., Zare Jeddi, M., Djidrovski, I., Friedrichs, S., Hardy, B., Hartung, T., Holland, D., Karwath, A., Kienhuis, A., Kleinstreuer, N., Lin, Z., Marczylo, E.L., Marvuglia, A., Qian, H., van Ravenzwaay, B., Rees, P., Sarimveis, H., Tralau, T., Wilmot, L., Zalewski, A., Rouquié, D., 2026. Building trust in the integration of artificial intelligence into chemical risk assessment: findings from the 2024 ECETOC workshop. Arch. Toxicol. 100 (5), 2149–2167. https://doi.org/10.1007/s00204-025-04286-8.

Gao, Q., Li, T., Zhu, G., Wang, J., Qiu, K., Liu, L., Yang, X., Wang, T., 2025. Estimating the causal effects of exposure mixtures: a generalized propensity score method. BMC Med. Res. Methodol. 25 (1), 221. https://doi.org/10.1186/s12874-025-02673-4.

García de Lomana, M., Marin Zapata, P.A., Montanari, F., 2023. Predicting the mitochondrial toxicity of small molecules: insights from mechanistic assays and Cell Painting data. Chem. Res. Toxicol. 36 (7), 1107–1120. https://doi.org/10.1021/acs.chemrestox.3c00086.

Giroux, E., 2025. Environmental health (the exposome). In: Schramme, T., Walker, M.J. (Eds.), Handbook of the philosophy of medicine. Springer, Dordrecht. https://doi.org/10.1007/978-94-024-2252-8_100.

Gonnabathula, P., Li, M., Nagumalli, S.K., Mehta, D., Fairman, K., 2025. Applications of PBPK modeling to estimate drug metabolism and related ADME processes in specific populations. Pharmaceutics 17 (9), 1207. https://doi.org/10.3390/pharmaceutics17091207.

Grzegorzewski, J., Brandhorst, J., König, M., 2022. Physiologically based pharmacokinetic (PBPK) modeling of the role of CYP2D6 polymorphism for metabolic phenotyping with dextromethorphan. Front. Pharmacol. 13, 1029073. https://doi.org/10.3389/fphar.2022.1029073.

Guimbaud, J.B., Siskos, A.P., Sakhi, A.K., Heude, B., Sabidó, E., Borràs, E., Keun, H., Wright, J., Julvez, J., Urquiza, J., Gützkow, K.B., Chatzi, L., Casas, M., Bustamante, M., Nieuwenhuijsen, M., Vrijheid, M., López-Vicente, M., de Castro Pascual, M., Stratakis, N., Robinson, O., Grazuleviciene, R., Slama, R., Alemany, S., Basagaña, X., Plantevit, M., Cazabet, R., Maitre, L., 2024. Machine learning-based health environmental-clinical risk scores in European children. Commun. Med. (Lond.) 4 (1), 98. https://doi.org/10.1038/s43856-024-00513-y.

Gustavsson, M., Molander, S., Backhaus, T., Kristiansson, E., 2023. Risk assessment of chemicals and their mixtures are hindered by scarcity and inconsistencies between different environmental exposure limits. Environ. Res. 225, 115372. https://doi.org/10.1016/j.envres.2023.115372.

Hartung, T., Kleinstreuer, N., 2025. Challenges and opportunities for validation of AI-based new approach methods. ALTEX 42 (1), 3–21. https://doi.org/10.14573/altex.2412291.

Herman, D., Kańduła, M.M., Freitas, L.G.A., van Dongen, C., Le Van, T., Mesens, N., Jaensch, S., Gustin, E., Micholt, L., Lardeau, C.H., Varsakelis, C., Reumers, J., Zoffmann, S., Will, Y., Peeters, P.J., Ceulemans, H., 2023. Leveraging Cell Painting images to expand the applicability domain and actively improve deep learning quantitative structure-activity relationship models. Chem. Res. Toxicol. 36 (7), 1028–1036. https://doi.org/10.1021/acs.chemrestox.2c00404.

Hou, M., Tang, S., Zhang, F., Fu, S., Ding, H., Cha, Y., Ma, X., Shi, Y., Cai, Y., 2025. Chemical exposure in females of childbearing age associated with sex hormones: evidence from an untargeted exposomic approach. Environ. Int. 197, 109362. https://doi.org/10.1016/j.envint.2025.109362.

Italiya, G., Subramanian, S., 2025. Leveraging new approach methodologies: ecotoxicological modelling of endocrine disrupting chemicals to Danio rerio through machine learning and toxicity studies. Toxicol. Mech. Methods 35 (2), 197–213. https://doi.org/10.1080/15376516.2024.2400324.

Jeong, J., Choi, J., 2022. Artificial intelligence-based toxicity prediction of environmental chemicals: future directions for chemical management applications. Environ. Sci. Technol. 56 (12), 7532–7543. https://doi.org/10.1021/acs.est.1c07413.

Jiang, J., Li, Y., Cao, S., Shan, Y., Liu, Y., Fei, T., Yu, Y., Feng, Y., Li, Y., Li, Y., Yuan, J., 2025. Artificial intelligence in bioinformatics: a survey. Brief. Bioinform 26 (6), bbaf576. https://doi.org/10.1093/bib/bbaf576.

Joubert, B.R., Kioumourtzoglou, M.A., Chamberlain, T., Chen, H.Y., Gennings, C., Turyk, M.E., Miranda, M.L., Webster, T.F., Ensor, K.B., Dunson, D.B., Coull, B.A., 2022. Powering research through innovative methods for mixtures in epidemiology (PRIME) program: novel and expanded statistical methods. Int. J. Environ. Res. Public. Health 19 (3), 1378. https://doi.org/10.3390/ijerph19031378.

Jusko, N.M., Karekar, N., Cao, A., Sun, D., 2026. AI and machine learning for predicting in vitro/preclinical drug toxicities and adverse events in clinical patients (in press.). Acta Pharm. Sin. B. https://doi.org/10.1016/j.apsb.2026.05.028.

Kar, S., Leszczynski, J., 2019. Exploration of computational approaches to predict the toxicity of chemical mixtures. Toxics 7 (1), 15. https://doi.org/10.3390/toxics7010015.

Karasaki, S., Morello-Frosch, R., Callaway, D., 2024. Machine learning for environmental justice: dissecting an algorithmic approach to predict drinking water quality in California. Sci. Total. Environ. 954, 175730. https://doi.org/10.1016/j.scitotenv.2024.175730.

Keil, A.P., Buckley, J.P., O'Brien, K.M., Ferguson, K.K., Zhao, S., White, A.J., 2020. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ. Health Perspect. 128 (4), 47004. https://doi.org/10.1289/EHP5838.

Ketkar, R., Liu, Y., Wang, H., Tian, H., 2023. A benchmark study of graph models for molecular acute toxicity prediction. Int. J. Mol. Sci. 24 (15), 11966. https://doi.org/10.3390/ijms241511966.

Kienzler, A., Bopp, S.K., van der Linden, S., Berggren, E., Worth, A., 2016. Regulatory assessment of chemical mixtures: requirements, current approaches and future perspectives. Regul. Toxicol. Pharmacol. 80, 321–334. https://doi.org/10.1016/j.yrtph.2016.05.020.

Kleinstreuer, N., Hartung, T., 2024. Artificial intelligence (AI) - it's the end of the tox as we know it (and I feel fine). Arch. Toxicol. 98 (3), 735–754. https://doi.org/10.1007/s00204-023-03666-2.

Kolesnyk, S., Souza, J.P., Bearth, A., Aicher, L., Wilks, M.F., Mosig, A., Fritsche, E., 2026. Bridging AI advancements with risk assessment needs: a journey towards effective use and regulatory acceptance. Toxicology 523, 154439. https://doi.org/10.1016/j.tox.2026.154439.

Komura, T., Bargagli-Stoffi, F.J., Arah, O.A., Inoue, K., 2026. Estimating and discovering heterogeneous treatment effects using machine learning in epidemiological studies: a practical guide. Int. J. Epidemiol. 55 (3), dyag092. https://doi.org/10.1093/ije/dyag092.

Li, Z., Chen, X., Wen, H., Zhang, R.Q., Li, M., Zhang, X., Yin, H., Yang, Q., Lam, K.Y., Lio, P., Yiu, S.M., 2026. A systematic survey and benchmark of deep learning for molecular property prediction in the foundation model era. J. Chem. Theory Comput. 22 (10), 4866–4887. https://doi.org/10.1021/acs.jctc.5c02081.

Limbu, S., Dakshanamurthy, S., 2023. Predicting dose-dependent carcinogenicity of chemical mixtures using a novel hybrid neural network framework and mathematical approach. Toxics 11 (7), 605. https://doi.org/10.3390/toxics11070605.

Lin, Z., Chou, W.C., 2022. Machine learning and artificial intelligence in toxicological sciences. Toxicol. Sci. 189 (1), 7–19. https://doi.org/10.1093/toxsci/kfac075.

Luechtefeld, T., Hartung, T., 2025. Navigating the AI frontier in toxicology: trends, trust, and transformation. Curr. Environ. Health Rep. 12 (1), 51. https://doi.org/10.1007/s40572-025-00514-6.

Luijten, M., Vlaanderen, J., Kortenkamp, A., Antignac, J.P., Barouki, R., Bil, W., van den Brand, A., den Braver-Sewradj, S., van Klaveren, J., Mengelers, M., Ottenbros, I., Rantakokko, P., Kolossa-Gehring, M., Lebret, E., 2023. Mixture risk assessment and human biomonitoring: lessons learnt from HBM4EU. Int. J. Hyg. Environ. Health 249, 114135. https://doi.org/10.1016/j.ijheh.2023.114135.

Madan, S., Lentzen, M., Brandt, J., Rueckert, D., Hofmann-Apitius, M., Fröhlich, H., 2024. Transformer models in biomedicine. BMC Med. Inform. Decis. Mak. 24 (1), 214. https://doi.org/10.1186/s12911-024-02600-5.

Malarvanann, M., Monohar, S., Kate, S.S., Taneja, I., Jaiswal, S., Pratap, B., Rathi, P.C., Thakur, S., Paul, D., Wahajuddin, M., 2026. A review on integrated machine learning and deep learning driven artificial intelligence models for pharmacokinetics and toxicokinetics predictions, and their application. Drug. Metab. Dispos. 54 (3), 100240. https://doi.org/10.1016/j.dmd.2026.100240.

Mao, J., Akhtar, J., Zhang, X., Sun, L., Guan, S., Li, X., Chen, G., Liu, J., Jeon, H.N., Kim, M.S., No, K.T., Wang, G., 2021. Comprehensive strategies of machine-learning-based quantitative structure-activity relationship models. iScience 24 (9), 103052. https://doi.org/10.1016/j.isci.2021.103052.

Moccia, C., Moirano, G., Popovic, M., Pizzi, C., Fariselli, P., Richiardi, L., Ekstrøm, C.T., Maule, M., 2024. Machine learning in causal inference for epidemiology. Eur. J. Epidemiol. 39 (10), 1097–1108. https://doi.org/10.1007/s10654-024-01173-x.

Mortensen, H.M., Gromelski, M., Hench, G., Martens, M., Wittwehr, C., Kumar, S., Kumar, V., Audouze, K., Virvilis, V., Nymark, P., Angrish, M., Lynch, I., Edwards, S., Magagna, B., Wojewodzic, M.W., 2025. FAIR AOP Cluster Working Group. The FAIR AOP roadmap for 2025: advancing findability, accessibility, interoperability, and re-usability of adverse outcome pathways. Comput. Toxicol. 35, 100368. https://doi.org/10.1016/j.comtox.2025.100368.

Muir, D.C.G., Getzinger, G.J., McBride, M., Ferguson, P.L., 2023. How many chemicals in commerce have been analyzed in environmental media? A 50 year bibliometric analysis. Environ. Sci. Technol. 57 (25), 9119–9129. https://doi.org/10.1021/acs.est.2c09353.

Muratov, E.N., Bajorath, J., Sheridan, R.P., Tetko, I.V., Filimonov, D., Poroikov, V., Oprea, T.I., Baskin, I.I., Varnek, A., Roitberg, A., Isayev, O., Fourches, Curtarolo S., Cohen, D., Aspuru-Guzik, Y., Winkler, A., Agrafiotis, D.A., Cherkasov, D., Tropsha, A., 2020. A. QSAR without borders. Chem. Soc. Rev. 49 (11), 3525–3564. https://doi.org/10.1039/d0cs00098a.

Nikolopoulou, D., Ntzani, E., Kyriakopoulou, K., Anagnostopoulos, C., Machera, K., 2023. Priorities and challenges in methodology for human health risk assessment from combined exposure to multiple chemicals. Toxics 11 (5), 401. https://doi.org/10.3390/toxics11050401.

Ortega-Vallbona, R., Palomino-Schätzlein, M., Tolosa, L., Benfenati, E., Ecker, G.F., Gozalbes, R., Serrano-Candelas, E., 2024. Computational strategies for assessing adverse outcome pathways: hepatic steatosis as a case study. Int. J. Mol. Sci. 25 (20), 11154. https://doi.org/10.3390/ijms252011154.

Pan, S., Li, Z., Rubbo, B., Quon-Chow, V., Chen, J.C., Baumert, B.O., Garcia, E., Aung, M. T., Conti, D.V., Chatzi, L., 2025. Applications of mixture methods in epidemiological studies investigating the health impact of persistent organic pollutants exposures: a scoping review. J. Expo. Sci. Environ. Epidemiol. 35 (4), 522–534. https://doi.org/10.1038/s41370-024-00717-3.

Pérez Santín, E., Rodríguez Solana, R., González García, M., García Suárez, M.D.M., Blanco Díaz, G.D., Cima Cabal, M.D., Moreno Rojas, J.M., López Sánchez, J.I., 2021. Toxicity prediction based on artificial intelligence: a multidisciplinary overview. WIREs Comput. Mol. Sci. 11 (5), e1516. https://doi.org/10.1002/wcms.1516.

Pressman, P., Hayes, A.W., 2026. Artificial intelligence-augmented mixture toxicology: reframing unresolved risk at Camp Lejeune. Toxicol. Sci. 209 (6), kfag066. https://doi.org/10.1093/toxsci/kfag066.

Reale, E., Zare Jeddi, M., Paini, A., Connolly, A., Duca, R., Cubadda, F., Benfenati, E., Bessems, J., Galea, K.S., Dirven, H., Santonen, T., Koch, H.M., Jones, K., Sams, C., Viegas, S., Kyriaki, M., Campisi, L., David, A., Antignac, J.P., Hopf, N.B., 2024. Human biomonitoring and toxicokinetics as key building blocks for next generation risk assessment. Environ. Int. 184, 108474. https://doi.org/10.1016/j.envint.2024.108474.

Ren, M., Wu, T., Zhang, H., Yang, S., Zhao, L., Zhuang, L., Lu, Q., Han, X., Pan, B., Li, T., Xue, J., Chen, Y., Bloom, M.S., Fang, M., Wang, B., 2025. Using environmental mixture exposure-triggered biological knowledge-driven machine learning to predict early pregnancy loss. Environ. Sci. Technol. 59 (37), 19691–19704. https://doi.org/10.1021/acs.est.5c05389.

Rider, C.V., 2022. Mixture math: deciding what to add in a cumulative risk assessment. Curr. Opin. Toxicol. 31, 100358. https://doi.org/10.1016/j.cotox.2022.100358.

Rogers, J.D., Bundy, J.L., Harrill, J.A., Judson, R.J., Paul-Friedman, K., Everett, L.J., 2025. Integrating transcriptomic and targeted new approach methodologies into a tiered framework for chemical bioactivity screening. Environ. Health Perspect. 133 (6), 67013. https://doi.org/10.1289/EHP16024.

Romero-Becerra, R., Zhao, Z., Nebdal, D., Müller, E., Bergholtz, H., Norum, J.H., Aittokallio, T., 2025. Improved analysis of in vivo drug combination experiments with a comprehensive statistical framework and web-tool. Nat. Commun. 16 (1), 10151. https://doi.org/10.1038/s41467-025-65218-9.

Santhappan, J.S., Ahmed Al-Shahri, A.S., Kalaiselvan, N., Baishya, N., Mathimani, T., 2026. A review on chemometrics for detecting emerging pollutants in the environment and assessing their toxicological impact on humans and the environment. iScience 29 (7), 116474. https://doi.org/10.1016/j.isci.2026.116474.

Schmitt, F.M., Viaene, K., Vlaeminck, K., Nys, C., Du Laing, G., De Schamphelaere, K., 2026 Jun 15. Chronic toxicity of metal-organic mixtures to Daphnia magna: To what extent can Concentration Addition and Independent Action predict effects? Environ. Toxicol. Chem., vgag163 https://doi.org/10.1093/etojnl/vgag163.

Seal, S., Mahale, M., García-Ortegón, M., Joshi, C.K., Hosseini-Gerami, L., Beatson, A., Greenig, M., Shekhar, M., Patra, A., Weis, C., Mehrjou, A., Badré, A., Paisley, B., Lowe, R., Singh, S., Shah, F., Johannesson, B., Williams, D., Rouquie, D., Clevert, D. A., Schwab, P., Richmond, N., Nicolaou, C.A., Gonzalez, R.J., Naven, R., Schramm, C., Vidler, L.R., Mansouri, K., Walters, W.P., Wilk, D.D., Spjuth, O., Carpenter, A.E., Bender, A., 2025. Machine learning for toxicity prediction using chemical structures: pillars for success in the real world. Chem. Res. Toxicol. 38 (5), 759–807. https://doi.org/10.1021/acs.chemrestox.5c00033.

Serra, A., Torres Maia, M., Tsiros, P., Minadakis, V., Riudavets-Puig, R., Perello-Y-Bestard, A., Nikiforou, F., Karakoltzidis, A., Di Lieto, E., Schaffert, A., Al-Abdulraheem, Z., Virmani, I., Dziubaniuk, O., Karhukorpi, S., Dokler, J., Zouraris, D., Mintis, D.G., Varsou, D.D., Tsoumanis, A., Melagraki, G., Isigonis, P., Papadiamantis, A.G., Buljan, M., Agalliadou, A., Ellis, L.A., Sergent, J.A., Alves Siqueira de Assunção, M., Martinez, D.S.T., Winkler, D., Park, S.G., Ha, S.M., Gerelkhuu, Z., Yoon, T.H., Karakitsios, S., Sarigiannis, D.A., Afantitis, A., Cucurachi, S., Serchi, T., Marvuglia, A., Exner, T., Siltaloppi, J., Paparella, M., Peijnenburg, W., Wick, P., Lynch, I., Sarimveis, H., Greco, D., 2026. Review on predictive models and integration strategies for holistic impact assessment of chemicals and materials. Environ. Sci. Technol. 60 (5), 3739–3766. https://doi.org/10.1021/acs.est.5c04489.

Sillé, F.C.M., Prasse, C., Luechtefeld, T., Hartung, T., 2025. AI redefines mass spectrometry chemicals identification: retention time prediction in metabolomics and for a Human Exposome Project. Front. Public. Health 13, 1687056. https://doi.org/10.3389/fpubh.2025.1687056.

Singh, A.V., Bhardwaj, P., Laux, P., Pradeep, P., Busse, M., Luch, A., Hirose, A., Osgood, C.J., Stacey, M.W., 2024. AI and ML-based risk assessment of chemicals: predicting carcinogenic risk from chemical-induced genomic instability. Front. Toxicol. 6, 1461587. https://doi.org/10.3389/ftox.2024.1461587.

Siroux, V., Agier, L., Slama, R., 2016. The exposome concept: a challenge and a potential driver for environmental health research. Eur. Respir. Rev. 25 (140), 124–129. https://doi.org/10.1183/16000617.0034-2016.

Soulios, K., Drosatos, G., Kaldoudi, E., 2023. deepFPlearn+: enhancing toxicity prediction across the chemical universe using graph neural networks. Bioinformatics 39 (12), btad713. https://doi.org/10.1093/bioinformatics/btad713.

Sun, X., Chu, J., Ni, R., Liu, F., Zhang, Y., Li, X., 2026. ChronicDPipredictor: an interpretable deep learning framework for chemical chronic and subchronic toxicity assessment. Mol. Divers. 30 (2), 2663–2676. https://doi.org/10.1007/s11030-026-11482-w.

Talavera Andújar, B., Schymanski, E.L., 2026. High resolution mass spectrometry to investigate the human exposome: where are we? Exposome 6 (1), osaf017. https://doi.org/10.1093/exposome/osaf017.

Thienpont, A., Verhulst, S., Van Grunsven, L.A., Rogiers, V., Vanhaecke, T., Mertens, B., 2023. Novel prediction models for genotoxicity based on biomarker genes in human HepaRG cells. ALTEX 40 (2), 271–286. https://doi.org/10.14573/altex.2206201.

Vermeulen, R., Schymanski, E.L., Barabási, A.L., Miller, G.W., 2020. The exposome and health: where chemistry meets biology. Science 367 (6476), 392–396. https://doi.org/10.1126/science.aay3164.

Vignaux, P.A., Lane, T.R., Urbina, F., Gerlach, J., Puhl, A.C., Snyder, S.H., Ekins, S., 2023. Validation of acetylcholinesterase inhibition machine learning models for multiple species. Chem. Res. Toxicol. 36 (2), 188–201. https://doi.org/10.1021/acs.chemrestox.2c00283.

Vijayaraghavan, S., Lakshminarayanan, A., Bhargava, N., Ravichandran, J., Vivek-Ananth, R.P., Samal, A., 2024. Machine learning models for prediction of xenobiotic chemicals with high propensity to transfer into human milk. ACS Omega 9 (11), 13006–13016. https://doi.org/10.1021/acsomega.3c09392.

Wambaugh, J.F., Rager, J.E., 2022. Exposure forecasting - ExpoCast - for data-poor chemicals in commerce and the environment. J. Expo. Sci. Environ. Epidemiol. 32 (6), 783–793. https://doi.org/10.1038/s41370-022-00492-z.

Wan, M., Simonin, E.M., Johnson, M.M., Zhang, X., Lin, X., Gao, P., Patel, C.J., Yousuf, A., Snyder, M.P., Hong, X., Wang, X., Sampath, V., Nadeau, K.C., 2025. Exposomics: a review of methodologies, applications, and future directions in molecular medicine. EMBO Mol. Med. 17 (4), 599–608. https://doi.org/10.1038/s44321-025-00191-w.

Wang, Z., Walker, G.W., Muir, D.C.G., Nagatani-Yoshida, K., 2020. Toward a global understanding of chemical pollution: a first comprehensive analysis of national and regional chemical inventories. Environ. Sci. Technol. 54 (5), 2575–2584. https://doi.org/10.1021/acs.est.9b06379.

Weichert, F.G., Inostroza, P.A., Ahlheim, J., Backhaus, T., Brack, W., Brauns, M., Fink, P., Krauss, M., Svedberg, P., Hollert, H., 2025. AI-aided chronic mixture risk assessment along a small European river reveals multiple sites at risk and pharmaceuticals being the main risk drivers. Environ. Int. 197, 109370. https://doi.org/10.1016/j.envint.2025.109370.

Wen, T., Cai, X., Li, J., 2025. Graph neural networks vs. traditional QSAR: a comprehensive comparison for multi-label molecular odor prediction. Molecules 30 (23), 4605. https://doi.org/10.3390/molecules30234605.

Wild, C.P., 2005. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol. Biomark. Prev. 14 (8), 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456.

Wild, C.P., 2012. The exposome: from concept to utility. Int. J. Epidemiol. 41 (1), 24–32. https://doi.org/10.1093/ije/dyr236.

Wu, Z., Wang, J., Du, H., Jiang, D., Kang, Y., Li, D., Pan, P., Deng, Y., Cao, D., Hsieh, C.Y., Hou, T., 2023. Chemistry-intuitive explanation of graph neural networks for molecular property prediction with substructure masking. Nat. Commun. 14 (1), 2585. https://doi.org/10.1038/s41467-023-38192-3.

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C.W., Qiu, J., Hua, K., Su, W., Wu, J., Xu, H., Han, Y., Fu, C., Yin, Z., Liu, M., Roepman, R., Dietmann, S., Virta, M., Kengara, F., Zhang, Z., Zhang, L., Zhao, T., Dai, J., Yang, J., Lan, L., Luo, M., Liu, Z., An, T., Zhang, B., He, X., Cong, S., Liu, X., Zhang, W., Lewis, J.P., Tiedje, J.M., Wang, Q., An, Z., Wang, F., Zhang, L., Huang, T., Lu, C., Cai, Z., Wang, F., Zhang, J., 2021. Artificial intelligence: a powerful paradigm for scientific research. Innov. (Camb.) 2 (4), 100179. https://doi.org/10.1016/j.xinn.2021.100179.

Yang, H., Sun, L., Li, W., Liu, G., Tang, Y., 2018. In silico prediction of chemical toxicity for drug design using machine learning methods and structural alerts. Front. Chem. 6, 30. https://doi.org/10.3389/fchem.2018.00030.

Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., Guzman-Perez, A., Hopper, T., Kelley, B., Mathea, M., Palmer, A., Settels, V., Jaakkola, T., Jensen, K., Barzilay, R., 2019. Analyzing learned molecular representations for property prediction. J. Chem. Inf. Model. 59 (8), 3370–3388. https://doi.org/10.1021/acs.jcim.9b00237.

Yu, L., Liu, W., Wang, X., Ye, Z., Tan, Q., Qiu, W., Nie, X., Li, M., Wang, B., Chen, W., 2022. A review of practical statistical methods used in epidemiological studies to estimate the health effects of multi-pollutant mixture. Environ. Pollut. 306, 119356. https://doi.org/10.1016/j.envpol.2022.119356.

Zare Jeddi, M., Galea, K.S., Viegas, S., Fantke, P., Louro, H., Theunis, J., Govarts, E., Denys, S., Fillol, C., Rambaud, L., Kolossa-Gehring, M., Santonen, T., van der Voet, H., Ghosh, M., Costa, C., Teixeira, J.P., Verhagen, H., Duca, R.C., Van

Nieuwenhuyse, A., Jones, K., Sams, C., Sepai, O., Tranfo, G., Bakker, M., Palmen, N., van Klaveren, J., Scheepers, P.T.J., Paini, A., Canova, C., von Goetz, N., Katsonouri, A., Karakitsios, S., Sarigiannis, D.A., Bessems, J., Machera, K., Harrad, S., Hopf, N.B., 2023. FAIR environmental and health registry (FAIREHR) - supporting the science to policy interface and life science research, development and innovation. Front. Toxicol. 5, 1116707. https://doi.org/10.3389/ftox.2023.1116707.

Zhang, F., Wang, Z., Peijnenburg, W.J.G.M., Vijver, M.G., 2023. Machine learning-driven QSAR models for predicting the mixture toxicity of nanoparticles. Environ. Int. 177, 108025. https://doi.org/10.1016/j.envint.2023.108025.

Zhang, J., Hu, Y., Wang, X., Ding, X., Cen, X., Wang, B., Yang, S., Ye, Z., Qiu, W., Chen, W., Zhou, M., 2024. Associations of personal PM2.5-bound heavy metals and heavy metal mixture with lung function: results from a panel study in Chinese urban residents. Chemosphere 364, 143084. https://doi.org/10.1016/j.chemosphere.2024.143084.

Zhang, J., Li, H., Zhang, Y., Huang, J., Ren, L., Zhang, C., Zou, Q., Zhang, Y., 2025. Computational toxicology in drug discovery: applications of artificial intelligence in ADMET and toxicity prediction. Brief. Bioinform 26 (5), bbaf533. https://doi.org/10.1093/bib/bbaf533.

Zhu, G., Wen, Y., Cao, K., He, S., Wang, T., 2024. A review of common statistical methods for dealing with multiple pollutant mixtures and multiple exposures. Front. Public. Health 12, 1377685. https://doi.org/10.3389/fpubh.2024.1377685.