Review | Open Access

Journal of Environmental Exposure Assessment

Wu *et al. J. Environ. Expo. Assess.* **2026**, *5*, 32 **DOI:**10.20517/jeea.2026.34

Artificial intelligence for toxicokinetic parameterization in environmental exposure assessment

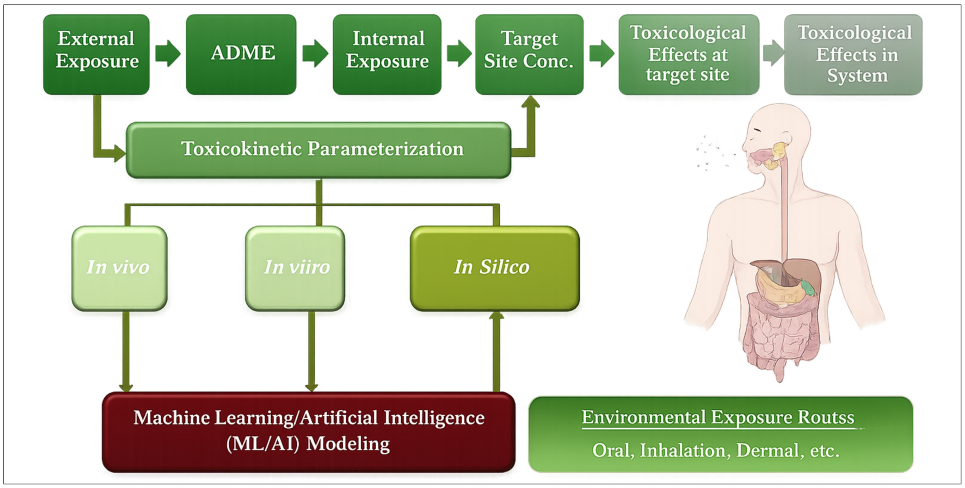
Xue Wu[1,2,3], Zhicheng Zhang[1,2,3], Kun Mi[1,2,3], Qiran Chen[4]

Keywords:
Toxicokinetics, exposure assessment, artificial intelligence, ADME

Citation: Wu, X.; Zhang, Z.; Mi, K.; Chen, Q. Artificial intelligence for toxicokinetic parameterization in environmental exposure assessment. *J. Environ. Expo. Assess.* **2026**, *5*, 32. https://dx.doi.org/10.20517/jeea.2026.34

Received: 16 Jun 2026
First Decision: 17 Jul 2026
Revised: 30 Aug 2026
Accepted: 8 Sep 2026
Published: 24 Sep 2026

Academic Editor:
Stuart Harrad
Copy Editor:
Pei-Yun Wang
Production Editor:
Pei-Yun Wang



Abstract

Toxicokinetics (TK) characterizes the absorption, distribution, metabolism, and excretion (ADME) of chemicals in the body and connects exposure with internal dose and potential toxicity. Because *in vivo* and *in vitro* TK assays are costly and low-throughput, and demand for emerging chemicals is growing, artificial intelligence and machine learning (AI/ML) offer alternative tools to predict ADME profiles for data-poor chemicals. This review examines recent developments in AI-based TK parameterization with relevance to environmental exposure assessment. First, we outline the essential elements of AI-based TK parameterization paradigms, including data sources, molecular representations, and commonly used ML algorithms. Second, the recent applications of AI/ML models across four ADME processes are introduced. Third, we discuss current limitations and future perspectives, including model reliability, interpretability, data availability for environmental

[1]Department of Environmental and Global Health, College of Public Health and Health Professions, University of Florida, Gainesville, FL 32611, USA.
[2]Center for Environmental and Human Toxicology, University of Florida, Gainesville, FL 32611, USA.
[3]Center for Pharmacometrics and Systems Pharmacology, University of Florida, Orlando, FL 32827, USA.
[4]Department of Health Toxicology, Key Laboratory for Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430000, Hubei, China.

Correspondence to: Dr. Qiran Chen, Department of Health Toxicology, Key Laboratory for Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430000, Hubei, China. E-mail: chenqiran@hust.edu.cn

chemicals, and integrating AI-predicted TK parameters into mechanistic models, such as physiologically based pharmacokinetic (PBPK) models and other downstream modeling frameworks. Although many existing studies rely on pharmaceutical or mixed chemical datasets, these approaches could support TK parameterization and exposure assessment for environmental chemicals. Further progress will depend on improved data quality, broader environmental chemical coverage, and more rigorous model validation and uncertainty evaluation.

INTRODUCTION

Quantitative exposure assessment of environmental chemicals relies on characterizing how exogenous substances are absorbed, distributed, metabolized, and excreted (ADME) within the human body[1,2]. Specific toxicokinetic (TK) parameters, such as the fraction unbound in plasma, intrinsic hepatic clearance, volume of distribution, oral bioavailability, tissue-to-plasma partition coefficients, and renal clearance, collectively determine the internal dose at target tissues and consequent toxicity[3]. TK parameters thus serve as a critical link between environmental exposure estimates and biologically relevant internal doses.

Conventionally, determining these TK parameters relies on a suite of costly and time-intensive *in vivo* or *in vitro* assays, which typically push per-chemical costs into the hundreds of thousands or millions of dollars, with timelines spanning months to years[4]. Compounding this challenge is the continuously expanding chemical landscape, which makes the existing data gap even more evident. The US EPA's Toxic Substances Control Act chemical inventory lists approximately 86,000 substances in commerce, yet fewer than a few hundred possess comprehensive *in vivo* TK datasets in humans[5-7]. While the ToxCast and Tox21 libraries collectively contain 9,404 unique chemicals, human *in vivo* and *in vitro* TK data are only available for approximately 7% and 11% of these chemicals, respectively[8]. This critical data gap hampers the identification of hazardous substances and the characterization of complex exposure patterns, both key challenges in modern environmental exposure science.

To fill these data gaps, researchers have turned to computational methods to estimate TK parameters across large numbers of chemicals. Early efforts centered on simple quantitative structure-activity relationship (QSAR) paradigms that relied on manually selected molecular features (such as octanol-water partition coefficients and hydrogen bond donor/acceptor counts) coupled with linear statistical methods, such as linear regression, partial least squares, and logistic regression. Although these models provided valuable estimates for specific endpoints, they were constrained by a narrow application scope, high reliance on linear assumptions, and poor generalization to new chemical classes. Machine learning (ML) methodologies move beyond traditional linear QSAR and capture complex relationships between chemical structures and biological endpoints[9]. By introducing different input features, such as physicochemical descriptors, ML models achieved marked improvements in predictive performance and interpretability. The emergence of deep learning (DL) architectures further advanced this field by enabling end-to-end learning of molecular representations directly from raw data[10]. These advancements signal a transformative expansion of the methodological toolkit available for *in silico* TK parameter estimation.

Despite this rapid methodological progress, persistent challenges remain in domain transferability, data availability, model interpretability, and real-world applicability. This review provides a comprehensive assessment of artificial intelligence and machine learning (AI/ML) approaches for predicting TK parameters for environmental chemicals, including industrial chemicals, pesticides, consumer product constituents, and emerging contaminants. Literature searches were conducted in PubMed (https://pubmed.ncbi.nlm.nih.gov/) and Google Scholar (https://scholar.google.com/) between December 2025 and May 2026, with no restriction on publication year. The literature search was last updated in September 2026. Three thematic term groups were used: (1) ("artificial intelligence" OR "machine learning"); (2) ("toxicokinetics" OR "pharmacokinetics"

OR "ADME" OR "physiologically based pharmacokinetic modeling"); and (3) individual toxicokinetic parameters, such as absorption rate, bioavailability, permeability, distribution, transplacental transfer efficiency (TTE), metabolism, and renal/biliary clearance. The search was conducted using the combination of group (1) AND group (2), or group (1) AND group (3). Reference lists of relevant reviews and primary studies were also screened to identify additional publications. Studies were included if they applied AI/ML approaches to predict quantitative or categorical ADME/TK endpoints, TK parameters, or concentration–time profiles. We prioritized studies involving environmental chemicals or demonstrating potential relevance to environmental exposure assessment. Studies focused exclusively on toxicity prediction without a direct connection to ADME or TK processes were excluded. To clarify the scope of AI-based TK parameterization, the endpoints reviewed here are distinguished by their relationship to quantitative TK modeling. These include direct TK parameters that can be incorporated into mechanistic models, quantitative surrogate endpoints that characterize specific kinetic processes but may require further translation before model implementation, and qualitative or categorical endpoints that provide supporting information for ADME characterization and TK model development. Direct prediction of concentration–time profiles may also be considered. First, it examines potential data sources, curated databases, and molecular input representations requisite for robust AI/ML model development. Second, it reviews current AI/ML methodologies and relevant examples across ADME processes, including direct TK parameters, quantitative surrogate endpoints, qualitative/supporting endpoints, and related concentration–time profile prediction applications. The value and representativeness of AI/ML in TK information should not be judged solely by improvements in predictive performance, but by its ability to generate biologically meaningful TK parameters that can support quantitative internal-dose reconstruction and mechanistic exposure modeling. Accordingly, we evaluate current AI/ML applications not only by model accuracy (ACC), but also by endpoint relevance, validation and chemical-space coverage, model reliability, and integration with TK mechanistic models. Finally, this review discusses some essential challenges and potential applications of these AI/ML model predictions. By synthesizing these domains, this review provides a clear roadmap for advancing AI/ML from an emerging computational tool to a robust, scientifically auditable foundation for modern environmental risk assessment.

GENERAL CONSIDERATIONS FOR AI-BASED ADME/TK MODELING

AI/ML modeling for TK parameterization generally follows a common workflow in which experimental or curated data are first assembled, translated into molecular representations and descriptors, and subsequently used for model development and evaluation [Figure 1][11]. Model performance depends not only on algorithm selection but also on data quality and comparability, the appropriateness of input representations, endpoint harmonization, and the rigor of model validation and applicability domain (AD) assessment[12]. This section summarizes these general methodological components underlying AI-based ADME/TK modeling.

Experimental and curated data sources

AI-based TK parameterization depends on high-quality data that integrate chemical structure, biological activity, and measured or inferred ADME/TK endpoints. The common principle of "garbage in, garbage out" is particularly relevant because prediction errors may arise not only from model limitations but also from noisy, limited, inconsistent, or poorly curated input data[13]. Compared with pharmaceutical ADME datasets, environmental TK datasets are often more heterogeneous because of the diverse chemical structures, population-specific TK profiles, and limited experimental coverage[14]. Inconsistent experimental designs, endpoint definitions, and study conditions can substantially reduce cross-study comparability.

TK data are typically obtained from *in vivo* experiments and supplemented by *in vitro* ADME assays[15]. *In vivo* studies provide time–concentration profiles in plasma, blood, urine, feces, and tissues following oral, intravenous, dermal, inhalation, or other exposure routes. These data can be used to estimate TK parameters

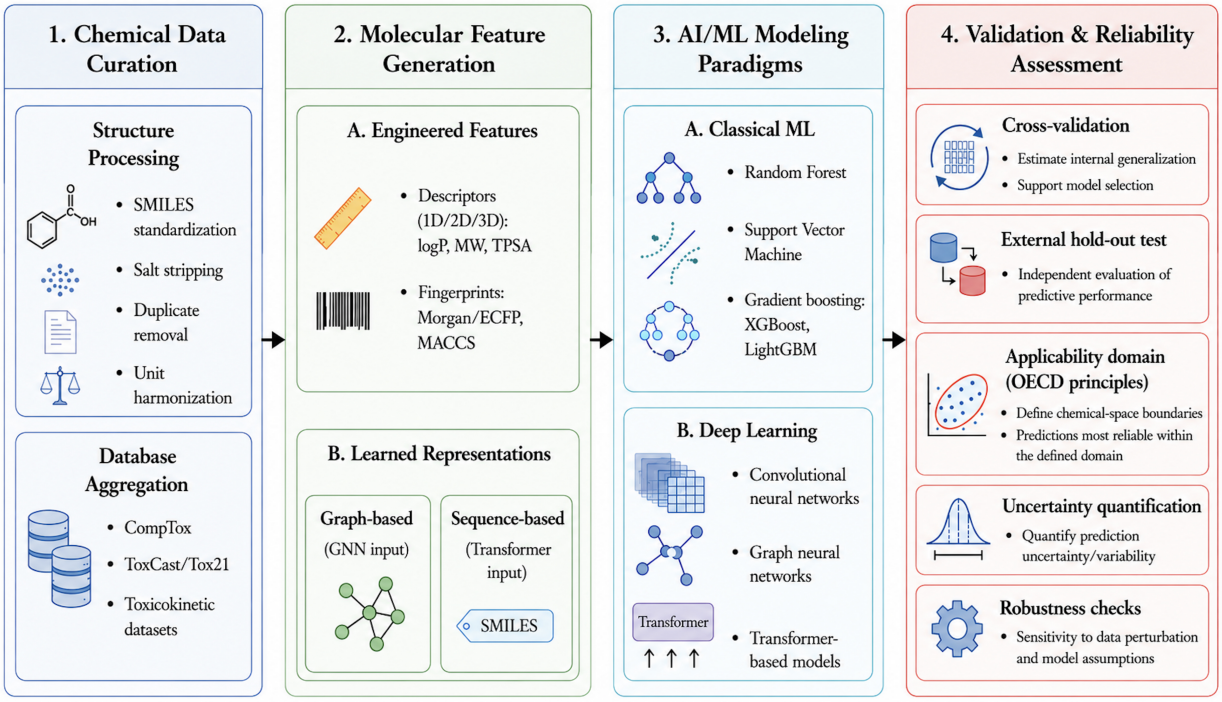


Figure 1. General workflow of AI/ML-based ADME prediction, including (1) chemical data curation; (2) molecular feature generation; (3) AI/ML model development; and (4) rigorous model validation [Created in BioRender. Zhang, Z. (2026) https://BioRender.com/86v7eo1]. AL/ML: Artificial intelligence and machine learning; ADME: absorption, distribution, metabolism, and excretion; SMILES: simplified molecular-input line-entry system; MW: molecular weight; TPSA: topological polar surface area; ECFP: extended-connectivity fingerprint; MACCS: molecular access system; GNN: graph neural network; OECD: Organisation for Economic Co-operation and Development.

such as clearance, half-life, bioavailability, volume of distribution, and tissue partition coefficient[16]. A major strength of *in vivo* TK studies is that they can reflect the integrated TK behavior of a given chemical in a biological system rather than considering specific TK characteristics separately[17]. However, as noted above, these datasets are often heterogeneous and limited for generalization because reported TK values may vary by species, sex, dose, sample size, analytical method, and model-fitting assumptions. For example, the mean half-life of perfluorooctanoic acid (PFOA) was reported to be approximately 0.15-0.19 days in female rats, while 1.6-18 days in males[18]. In practice, this variability may influence reported TK values as strongly as chemical structure itself[19]. Moreover, due to the interspecies ADME differences, the TK parameters from *in vivo* animal experiments may cause systematic bias when extrapolated for human exposure assessment. Therefore, when using TK parameters from *in vivo* experiments for AI/ML modeling, record these sources of variability as metadata for model development and interpretation.

Compared with *in vivo* experimental data, *in vitro* ADME assays facilitate standardization of experimental settings and support high-throughput chemical testing. Common endpoints of *in vitro* assays may include intestinal permeability, plasma protein binding, fraction unbound, intrinsic clearance in microsomes or hepatocytes, metabolic stability, and transporter-related activity[20]. Compared with animal-based *in vivo* tests for TK parameters, a key strength of *in vitro* assays is the adoption of human cells for testing, which can reduce interspecies TK variability and improve the reliability of results for human exposure assessment[21]. Moreover, as *in vitro* assays can quantify individual TK processes, they can support targeted screening of specific TK parameters across diverse chemicals, which greatly benefits AI/ML modeling and structural generalization for data-poor chemicals and chemical families, thereby accelerating exposure characterization of emerging environmental chemicals. Nevertheless, it needs to be careful when merging *in vivo* and *in vitro* TK data into a single database due to the two systems reflecting different biological contexts and exposure

conditions. In contrast to TK parameters derived from *in vivo* experiments, *in vitro* outcomes such as effective doses have to be extrapolated to *in vivo* equivalent doses via a reverse dosimetry approach, which introduces extra uncertainty and variability into the compiled database[6].

Beyond individual experiments, AI/ML-based TK parameterization generally relies on curated databases that integrate measurements across multiple compounds, studies, and experimental conditions. The quantity and quality of these data can substantially impact the prediction ACC and model robustness[22]. A representative example is the U.S. EPA's httk R package, which includes human *in vitro* TK parameters (plasma protein binding and hepatic clearance), structure-derived physicochemical properties, and species-specific physiological data[23]. Combined with a generic physiologically-based toxicokinetic (PBTK) model for *in vitro- in vivo* extrapolation, researchers can predict *in vivo* TK profiles or parameters [e.g., maximum concentration (C_{max}) and area under concentration curves over 24 h (AUC_{24h}), mean concentration (C_{mean})] for diverse chemicals in humans and animals[23,24]. Additional resources, such as PubChem and OPERA, can support chemical standardization, descriptor generation, and access to physicochemical and ADME-related properties[25,26].

Besides these general-purpose resources, class-specific databases have also been developed for emerging contaminants whose physicochemical and toxicokinetic characteristics may not be adequately represented by conventional chemical datasets. For example, microplastic/nanoparticle TK profiles can be influenced not only by chemical composition, but also by particle-specific properties, such as size, shape, and surface chemistry[27]. General-purpose chemical databases or molecular representations cannot adequately capture the particle-specific concentration-time curves and the determinants of ADME properties, thereby reducing the transferability and mechanistic relevance of AI-based TK parameterization for this class of microplastics. To address this limitation, a few databases tailored to microplastics/nanoparticles have been established in recent years, facilitating AI/ML development for TK estimation and improving explainability[28-30].

In practice, researchers often need to curate and integrate datasets from multiple studies to meet specific modeling objectives. Before data merging, chemical identifiers, endpoint definitions, measurement units, species, sex, life stage, exposure route, dose, sampling time, and other biological and experimental variables should be harmonized to ensure that records represent comparable biological and chemical quantities. Following harmonization, numerical features and continuous TK endpoints may require transformation and scaling because real-world datasets often contain variables with substantially different distributions and magnitudes. Without appropriate scaling, variables with larger numerical ranges may disproportionately influence model optimization and reduce prediction reliability, particularly for scale-sensitive algorithms, such as k-nearest neighbors (KNNs), support vector machines (SVMs), and neural networks. Common approaches include min-max normalization and z-score standardization[31]. Min–max normalization rescales values to a predefined range, whereas z-score standardization centers each feature around its mean and scales it by its standard deviation. Additional preprocessing strategies, including logarithmic transformation, robust scaling, and categorical variable encoding, may also be applied depending on the dataset characteristics and modeling objectives[31]. A proper scaling strategy can help ML algorithms converge faster during training, avoid bias toward large magnitude variables, improve numerical stability, and produce more reliable predictions[32].

Input representations and descriptors

The predictive performance of AI/ML models depends not only on data availability, but also on how chemical and biological information is represented for model development. In AI-based ADME/TK modeling, input representations convert chemicals into machine-readable formats, whereas molecular descriptors provide quantitative variables derived from these representations to encode structural and

Table 1. Input representations and descriptors used in AI-based TK modeling

Input type	Examples	TK/ADME relevance	Ref.
Molecular descriptors	Molecular weight, logP/logD, pKa, TPSA, H-bond donors/acceptors, rotatable bonds, charge	Related to permeability, solubility, protein binding, metabolic stability, and clearance	Handa *et al.*, 2025[35]
Molecular fingerprints	MACCS keys, Morgan/ECFP fingerprints	Encode structural fragments associated with ADME properties	Ryu *et al.*, 2023[36]
Graph-based features	Atom–bond graphs, GNN embeddings	Capture molecular connectivity and local chemical environments	Xu *et al.*, 2017 Jiang *et al.*, 2021[37,38]
Physicochemical properties	Solubility, ionization state, lipophilicity, vapor pressure, Henry's law constant	Influence absorption, distribution, partitioning, and exposure route-specific behavior	Noga and Jurowski, 2025[39]
Biological context variables	Species, tissue composition, plasma protein levels, enzyme/transporter expression, life stage, sex	Modify distribution, metabolism, clearance, and internal dose	Li *et al.*, 2024[40]
Experimental metadata	Dose, administration route, dosage form, sampling time, and experimental conditions	Explains variability in measured TK endpoints	Fuhrer *et al.*, 2024[41]

AI: Artificial intelligence; TK: toxicokinetic; ADME: absorption, distribution, metabolism, and excretion; logP: octanol–water partition coefficient; pKa: acid dissociation constant; TPSA: topological polar surface area; MACCS: molecular access system; ECFP: extended-connectivity fingerprint; GNN: graph neural network.

physicochemical properties[33]. The choice of input representation can substantially influence prediction performance and model transferability, as different molecular representations may capture distinct structural information and perform differently across datasets and endpoints[34]. Major input representations and descriptors are summarized in Table 1.

Traditional molecular descriptors remain the most widely used inputs for TK-related prediction. These descriptors typically include molecular weight, logP/logD, pKa, hydrogen bond donors and acceptors, topological polar surface area, rotatable bonds, formal charge, and solubility. Their main advantage is interpretability, since many of these variables are mechanistically related to the key ADME properties, such as membrane permeability, tissue binding, plasma protein binding, metabolic clearance, and excretion behavior[35]. In addition to descriptors, molecular fingerprints, such as molecular access system (MACCS) keys and extended-connectivity fingerprints (ECFP), are frequently used to encode structural fragments and substructures into a machine-readable form[42]. These representations are particularly useful for nonlinear machine-learning models because they capture local structural patterns efficiently, although they are generally less interpretable than descriptor-based inputs and may be more sensitive to chemical-domain mismatch.

More recently, graph-based representations have become attractive because they model molecules directly as atoms and bonds, without requiring predefined descriptors. In this framework, graph neural networks (GNNs) can learn structure-property relationships from molecular topology and local chemical environments, potentially capturing higher-order structural features relevant to ADME endpoints[43]. For example, Ng and Lu evaluated GNNs combined with transfer learning for oral bioavailability prediction, achieving a final average ACC of 0.797, an F1 score of 0.840, and an area under the receiver operating characteristic curve (AUC-ROC) of 0.867, which outperformed previous studies using traditional molecular descriptors and fingerprints with the same test dataset[44].

In addition, biological context variables are also important for TK prediction. Factors such as tissue composition, gastrointestinal physiology, renal function, sex, age, and life stage can strongly influence ADME behavior[21,45]. For this reason, hybrid input strategies that combine chemical descriptors with biological and experimental metadata are often more informative than structure-only representations,

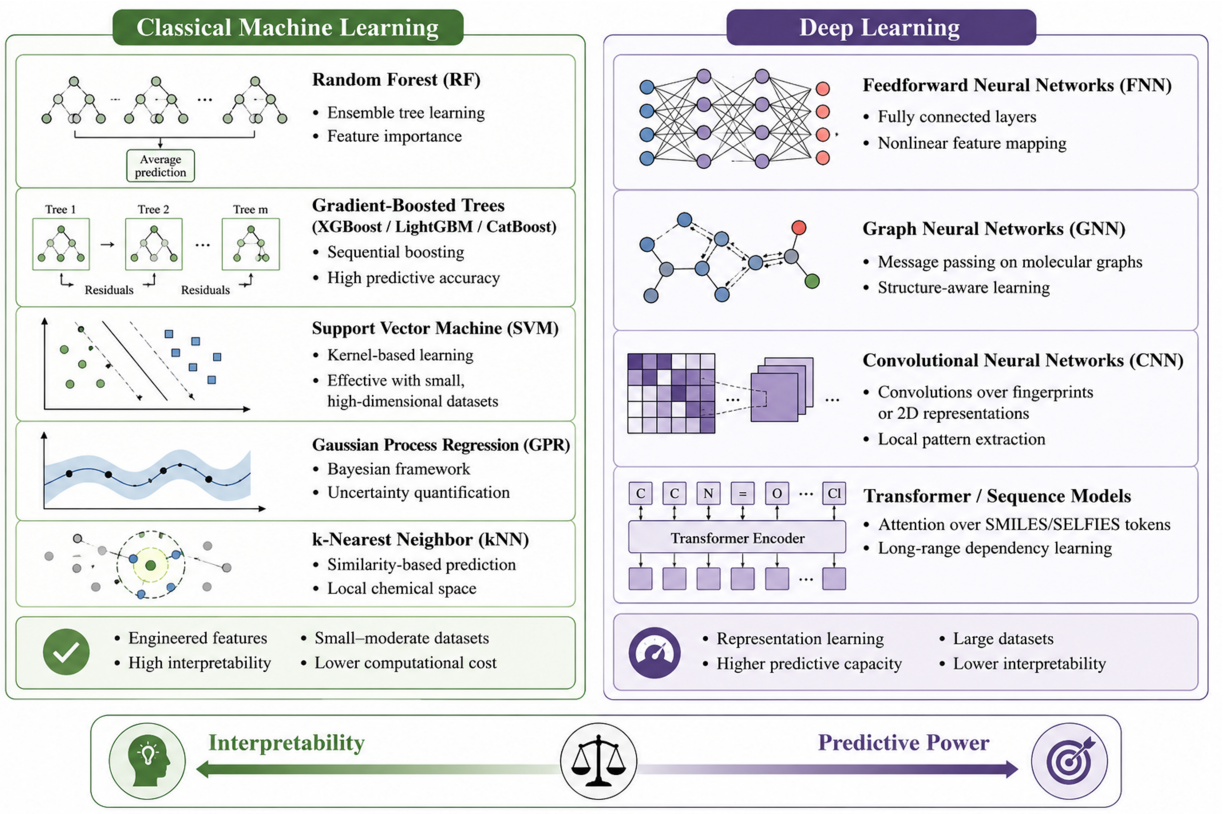


Figure 2. Representative AI/ML model architectures used for toxicokinetic parameterization [Created in BioRender. Zhang, Z. (2026) https://BioRender.com/mdf4f4m]. AI/ML: Artificial intelligence and machine learning; SMILES: simplified molecular-input line-entry system; SELFIES: self-referencing embedded strings.

particularly when the goal is to support physiologically based pharmacokinetic (PBPK) modeling and exposure assessment.

AI/ML modeling and evaluation

Following the translation of chemical structure and physicochemical features into numerical representations, AI/ML models can be developed to predict ADME- and TK-related endpoints. The current algorithmic landscape falls into two major paradigms: classical ML and DL. The representative model architectures are summarized in Figure 2.

Among classical ML methods, ensemble tree-based algorithms are widely applied in ADME/TK prediction. Random forests (RFs) are ensembles of decorrelated decision trees trained on bootstrap samples with random feature subsets. They offer robust performance, natural resistance to overfitting in high-dimensional spaces, and built-in feature importance estimation[46]. Gradient-boosted decision trees (e.g., XGBoost, LightGBM, and CatBoost) construct sequential ensembles where each successive tree corrects the residual errors of its predecessor[47]. These models typically achieve superior ACC on moderate-sized datasets, though they are highly sensitive to model tuning. SVMs with radial basis function or Tanimoto kernels remain competitive for regression on continuous TK endpoints, particularly when training sets are small and high-dimensional[48]. Gaussian process regression provides a principled Bayesian framework yielding both point predictions and predictive uncertainty estimates, though its cubic scaling with dataset size limits applicability to large training corpora[49]. The KNN algorithm is conceptually straightforward but foundational. It underpins the US EPA's OPERA suite, where it combines property forecasting with AD assessments based on local chemical similarity[50]. Overall, these classical algorithms pair naturally with the engineered feature

representations and remain the algorithms of choice when dataset sizes are modest (hundreds to low thousands of compounds) or when maximal interpretability is required.

The DL paradigm extends beyond classical methods by learning hierarchical feature representations directly from molecular inputs. Different neural-network structures can be paired with different molecular representations, such as fixed-length descriptors or fingerprints, molecular graphs, and sequence-based representations[33,34]. Early deep-learning applications to ADME/TK prediction commonly used feedforward neural networks (FNNs), including multilayer artificial neural networks (ANNs) and their deeper variants, deep neural networks (DNNs), which typically operate on fixed-length molecular descriptors or fingerprints[51]. These models retained the descriptor-based input pipeline of classical ML, but replaced linear or shallow nonlinear estimators with multilayer architectures capable of capturing higher-order interactions among descriptors. Convolutional neural networks (CNNs) use convolutional operations to learn local patterns from structured molecular encodings and can support automatic feature extraction[52]. Using graph-based descriptors as we discussed in the last section, GNN architectures aggregate local atomic environments into expressive molecular embeddings through iterative message passing and attention-weighted readout functions[53]. For sequence-based molecular representations, transformer-based models such as ChemBERTa and MoLFormer process sequence-based simplified molecular-input line-entry system (SMILES) inputs through multi-head self-attention layers, capturing long-range token dependencies that recurrent architectures handle less effectively[54,55]. Additionally, these models are not always applied independently and can be integrated into hybrid models to exploit complementary molecular representations and learning strategies. For example, Limbu *et al.* developed a hybrid CNN-FNN model (HNN) to predict the toxicity of thousands of chemicals[56]. The HNN model consisted of a CNN processing one-hot-encoded SMILES and an FNN processing molecular descriptors[56]. Such integration allows structure-derived features and engineered physicochemical descriptors to contribute jointly to prediction and may improve model robustness when information from one representation is limited.

The choice of modeling strategy should ultimately depend on dataset size, endpoint characteristics, and the required balance between ACC and interpretability. As schematically illustrated in Figure 2, simpler models are often favored when datasets are small or when greater transparency is required, whereas more flexible deep-learning models may improve predictive performance when sufficiently large and diverse datasets are available[57]. However, this increase in flexibility may also reduce interpretability and increase the risk of overfitting or poor extrapolation to novel chemical classes. Based on the task type (classification or regression), different metrics for model evaluation were used. Continuous endpoints, such as intrinsic clearance, fraction unbound, or volume of distribution, are commonly assessed using the coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE). Categorical endpoints, such as CYP inhibition or substrate classification, are typically evaluated using ACC, sensitivity, specificity, the AUC-ROC, and the Matthews correlation coefficient (MCC).

After training the models with an existing database and tuning it for good performance, the selected models still demand validation at different levels: internal cross-validation, external hold-out testing, and an AD assessment[58]. Internal validation through k-fold cross-validation (typically 5-fold or 10-fold) provides a baseline estimate of model performance but can yield optimistically biased results when applied to datasets with structural redundancy or activity[59,60]. External hold-out testing provides the most informative measure of true predictive ability. This process evaluates the trained model on an entirely independent dataset withheld from all stages of model development. Furthermore, AD assessment identifies the specific region of chemical space where a model's predictions are considered reliable. This assessment is a mandatory component of the Organisation for Economic Co-operation and Development (OECD) validation principles for regulatory QSAR models[61]. It is especially critical when applying models to the highly diverse structural

landscape of environmental chemicals. Common AD methods include leverage-based approaches (Williams plots), distance-to-model metrics in descriptor space, local density estimation, and conformal prediction frameworks that provide prediction-specific confidence levels[62]. The US EPA's OPERA suite exemplifies the integration of AD assessment into routine ADME prediction by providing per-chemical reliability indices alongside property estimates, enabling users to distinguish high-confidence predictions from extrapolations[50].

Finally, the selected model will be examined for mechanistic interpretation, uncertainty quantification, and downstream integration strategies with the aim of translating these validated ADME predictions into actionable exposure estimates.

APPLICATION OF AI MODELING FOR ADME PARAMETERIZATION

AI for absorption

Absorption is the process by which a chemical enters systemic circulation following external exposure. Environmental chemicals can be mainly absorbed through oral ingestion, inhalation, and dermal contact. Among these routes, oral absorption has been the most commonly studied in current AI/ML-based TK parameterization, largely because oral exposure is the most common route for conventional drugs and many tested chemicals, resulting in greater data availability and more established assay systems. Typically, oral absorption is characterized using parameters such as intestinal permeability, oral bioavailability, and the absorption rate constants. Oral bioavailability has received considerable attention, particularly from pharmaceutical research, as this parameter is a key determinant of the success or failure of candidate drugs. Multiple predictive approaches, including QSAR models, PBPK-based models, and more recent AI/ML algorithms, along with various software, have been developed for predicting oral bioavailability[63]. Similarly, researchers have developed a few models for other oral absorption-related parameters [Table 2]. For example, Ng and Lu evaluated a GNN combined with transfer learning to predict oral bioavailability. In their framework, the model was first pretrained on a solubility prediction task using graph-based molecular representations, and then fine-tuned for oral bioavailability classification. The best transfer-learning model achieved an ACC of 0.797, an F1 score of 0.840, and an AUC-ROC of 0.867, outperforming traditional ML models using descriptor- and fingerprint-based representations[44]. Shapley Additive exPlanations (SHAP) analysis further reported the most important factor in the quantitative estimation of drug-likeness (QED)[44]. This study illustrates the potential advantage of learned molecular representations and knowledge transfer for improving prediction when endpoint-specific datasets are limited. However, whether this improvement generalizes to structurally diverse environmental chemicals requires further validation because the AD analysis of the transfer model suggested the absence of distinctly out-of-domain compounds in the test set. In another instance, Kamiya *et al.* developed a lightGBM model to predict influx and efflux apparent permeability (P_{app}) across Caco-2 monolayers for 218 disparate chemicals by integrating *in vitro* permeability coefficients with 17 and 19 descriptors, respectively, such as logD, logP, topological polar surface area, molecular weight, and basic group count[65]. Their best lightGBM models achieved correlation coefficients of 0.83-0.84 for influx and efflux P_{app} (log-transformed, nm/s) prediction, indicating that AI/ML models can provide reasonably accurate estimates of intestinal permeability for structurally diverse chemicals.

Besides oral absorption, inhalation and dermal absorption have received less attention in current AI/ML-based parameterization due to limited availability of curated datasets. Nevertheless, these routes are highly relevant for environmental exposure assessment, especially for volatile chemicals, aerosols, consumer product ingredients, and occupational exposures. A few studies have begun to explore AI/ML approaches for inhalation- and dermal-related TK characterization using relatively small datasets. For example, Chiu *et al.* developed an Extra Trees-based multitask learning model to predict the isolated perfused lung absorption rate constant ($k_{a,IPL}$, min^{-1}) for a dataset of 29 chemicals. To address the limited pulmonary absorption data,

Table 2. Selected AI/ML studies to predict chemical absorption parameters

Parameter (transformation; unit; endpoint type)	Input features	Chemicals	Model	Results	Ref.
Oral					
HIA (classification: highly absorbed > 30% *vs.* poorly < 30%; unitless; qualitative/supporting endpoint)	10 optimal descriptors (selected from 1,529 features using DRAGON and TSAR)	844 highly absorbed (N_{train} = 489, N_{test} = 355), 398 poorly absorbed (N_{train} = 256, N_{test} = 142)[a]	SVM, ANN, KNN, PNN, PLS, and LDA	Best performance of SVM model (10-fold cross-validation ACC = 90.38%, ACC_{test} = 91.54%, ROC-AUC = 0.885)	Kumar *et al.*, 2017[64]
Oral bioavailability (classification: high ≥ 50%, low < 50%; unitless; qualitative/supporting endpoint)	♦ Pretraining: graph-based representations ♦ RF model: 45 molecular descriptors and fingerprints (Morgan FP, RDKit FP, MACCS keys) from chemical structures	Pretraining: 9,940 molecules Present study: 1,447 chemicals (N_{train} = 1,157, N_{test} = 290)[b]	GNN & Transfer model, RF	♦ RF with molecular descriptors had the best performance among all RF models ♦ SHAP analysis suggested that molecules with higher quantitative drug-likeness estimates exhibited higher oral bioavailability ♦ The transfer learning model showed the highest prediction performance among all models (ACC = 79.7%, AUC-ROC = 0.867)	Ng and Lu, 2023[44]
Intestinal permeability coefficient (log-transformed; nm/s; quantitative surrogate endpoint)	17 (A → B)/19 (B → A) descriptors, including logD, logP, MW, topological polar surface area, *etc.*	219 disparate chemicals[a]	Trivariate linear regression, LightGBM	Trivariate regression showed significant correlations between observed and predicted $logP_{app}$ using molecular weight and pH-dependent logD values (A → B: r = 0.76, n = 198; B → A: r = 0.77, n = 202). LightGBM further improved prediction ACC, reaching r = 0.83-0.84 (P < 0.001) for influx and efflux $logP_{app}$ prediction	Kamiya *et al.*, 2021[65]
Caco-2 permeability coefficient (log-transformed, cm/s; quantitative surrogate endpoint)	0-2D PaDEL descriptors (Morgan fingerprint, RDKit2D, molecular graphs)	5,654 compounds; an additional 67 compounds from Shanghai Qilu's *in-house* collection and 271 ChEMBL compounds for external validation[b]	RF, XGBoost, SVM, GBM, DMPNN, and CombinedNet	R^2 values ranged from ~0.4 to ~0.65 for the test set XGBoost model with combined molecular representations (all three types) had the best performance (R^2 = 0.622, RMSE = 0.487)	Wang *et al.*, 2025[66]
PAMPA permeability coefficient (log-transformed, cm/s; quantitative surrogate endpoint)	2,792 molecular descriptors	393 molecules (train: test ratio = 8:2)[b]	ANN, SVM	ANN model showed the best performance (R^2 = 0.84 for the external test set)	Racz *et al.*, 2023[67]
Dermal					
Partition coefficients for the stratum corneum and viable epidermis/dermis (K_{SC} and K_{VED}; log-transformed; unitless; direct TK parameter) Diffusion coefficients in two skin layers (D_{SC} and D_{VED}; log-transformed; cm^2/h; direct TK parameter)	Physicochemical descriptors, including molecular weight, lipophilicity, and HOMO/LUMO-related electronic descriptors	54 chemicals (N_{train} = 43, N_{test} = 11)[b]	Gradient boosting tree	Test-set prediction: R^2 = 0.776 ± 0.009 ($logK_{SC}$), 0.776 ± 0.019 ($logD_{SC}$), 0.912 ± 0.006 ($logK_{VED}$), and 0.754 ± 0.021 ($logD_{VED}$)	Narita *et al.*, 2025[68]
Dermal absorption (%; quantitative surrogate endpoint)	11 descriptors (logP, molecular weight, water-based formulation, dilution concentration, *etc.*)	ProHuma/ECPA dataset: 248 active substances from 25 formulation types at different concentrations[a]	Bayesian additive regression trees	Log Pow and molecular weight have the highest importance	Sarti *et al.*, 2025[69]

Inhalation					
Isolated perfused lung absorption rate constant (min^{-1}; direct TK parameter)	Seven selected molecular descriptors; shared permeability information from Caco-2 permeability, Calu-3 permeability, and isolated perfused lung absorption datasets	♦ IPL dataset: 29 chemicals with Ka_{IPL} data ♦ Caco-2 dataset: 960 chemicals with *in vitro* $Papp_{Caco-2}$ data ♦ Calu-3 dataset: 73 chemicals with *in vitro* $Papp_{calu3}$ data[b]	Extremely randomized trees and a multitask learning extension, MT-ExtraTrees, in the ExtraTrees package	The model achieved a correlation coefficient of r = 0.84 between predicted and observed Ka_{IPL} values in the independent test set	Chiu *et al.*, 2024[70]
Blood concentration–time profiles (concentration–time profile prediction)	Airborne VOC concentration, exposure time, initial blood concentration, and longitudinal blood concentration data	DBM and MCF exposure experiments[b]	Neural ODE	For DBM, Neural ODEs achieved a MAPE of 6.56% at 10,000 ppm, outperforming PBPK at low-to-moderate exposure levels. Neural ODEs yielded a MAPE of 25.55% for MCF at 10,000 ppm, though ACC declined at lower concentrations, such as 10 ppm	Simon, 2025[71]

[a]Studies used mixed environmental and pharmaceutical datasets (mixed relevance). [b]Studies are primarily based on pharmaceutical datasets and are considered indirectly relevant to environmental TK applications. AI/ML: Artificial intelligence and machine learning; HIA: human intestinal absorption; SVM: support vector machine; ANN: artificial neural network; KNN: k-nearest neighbor; PNN: probabilistic neural network; PLS: partial least squares; LDA: linear discriminant analysis; ACC: accuracy; ROC: receiver operating characteristic curve; AUC: area under the curve; RF: random forest; GNN: graph-based neural network; SHAP: Shapley Additive exPlanations; MW: molecular weight; XGBoost: extreme gradient boosting; GBM: gradient boosting machine; DMPNN: directed message passing neural network; RMSE: root mean squared error; PAMPA: parallel artificial membrane permeability assay; TK: toxicokinetics; HOMO: highest occupied molecular orbital; LUMO: lowest unoccupied molecular orbital; ECPA: European Crop Protection Association; IPL: isolated perfused lung; VOC: volatile organic compound; DBM: dibromomethane; MCF: methylchloroform; ODE: ordinary differential equation; MAPE: mean absolute percentage error; PBPK: physiologically-based pharmacokinetic.

the model jointly learned from the larger Caco-2 and Calu-3 permeability datasets and achieved a correlation coefficient of r = 0.84 between the predicted and observed $k_{a,\mathrm{IPL}}$ values in the independent test set[70]. Compared with single-task learning approaches, this study illustrates the advantages of multi-task learning to address limited endpoint-specific data availability by jointly learning from auxiliary tasks (i.e., Caco-2 and Calu-3 permeability in this study)[70]. The effectiveness of such multitask learning largely depends on the biological relevance and similarity of auxiliary tasks to the target task. In another instance, Simon applied a neural ordinary differential equation (ODE) model to predict blood concentration–time profiles of two volatile organic compounds (VOCs) following inhalation exposure[71]. In this approach, neural ODEs learn the chemical concentration changes by parameterizing the rate of change as a continuous function of time, providing a potential AI-based approach for environmental TK modeling, where chemical-specific kinetic parameters and large-scale training datasets are often limited.

For dermal exposure, Narita *et al.* developed a gradient boosting tree model to predict four skin permeation parameters: partition coefficients for the stratum corneum and viable epidermis/dermis (K_{SC} and K_{VED}) and diffusion coefficients in these two skin layers (D_{SC} and D_{VED})[68]. The predicted parameters were subsequently incorporated into a two-layer diffusion model to estimate finite-dose dermal permeation profiles. The model showed generally reasonable agreement with observed profiles for the evaluated chemicals. However, additional external validation is needed before broader application of this model in dermal pharmaceutical development and exposure assessment for topically exposed chemicals. In future applications, AI-based dermal absorption models could be used to evaluate the effectiveness of chemical neutralization and decontamination products, such as reactive skin decontamination lotion[72]. These models could incorporate decontamination efficiency, treatment timing, chemical reactivity, and residual dermal absorption to support exposure assessment following hazardous-material incidents.

AI for distribution

Following absorption through different exposure routes, chemicals are transported from the systemic circulation into tissues, organs, and biological barriers. In exposure assessment, this process is important because toxicity is often driven more directly by target-site concentrations than by external exposure dose alone.

AI modeling for distribution mainly focuses on predicting relevant parameters, including tissue: plasma partition coefficients, apparent volume of distribution, fraction unbound in plasma, and barrier permeability [Table 3]. For example, Dawson *et al.* developed QSAR models to predict *in vitro* fraction unbound in plasma for 2,057 compounds [1,308 pharmaceuticals and 749 compounds from the U.S. EPA Toxicity Forecaster (ToxCast) program][73]. The best model integrated four open-source chemical descriptors, including PaDEL, OPERA, ToxPrints, and MACCS, into an RF model and achieved an R^2 of 0.591 with an RMSE of 0.187[73]. This study is particularly relevant to environmental TK modeling because the predicted parameters were designed to support HTTK applications and internal dose estimation for data-poor chemicals. For highly protein-bound contaminants such as per- and polyfluoroalkyl substance (PFAS), uncertainty in predicting the unbound fraction affects the assessment of persistence and internal exposure.

Unlike general tissue partitioning, AI/ML models have also been applied to predict chemical transport across physiological barriers that regulate access to specific compartments. Blood–brain barrier (BBB) penetration is one of the most important specialized barriers, involving tight junctions, selective transport, and active efflux mechanisms. The BBB can prevent 98% of circulating molecules from entering the brain. Traditional experimental methods make it difficult to assess BBB penetration rates. AI/ML models can identify key features influencing BBB permeability, supporting TK characterization and exposure assessment. Early BBB-prediction models relied on traditional structural and physicochemical features, such as molecular weight, lipophilicity, and hydrogen-bonding properties[85]. For example, Liu *et al.* developed ML and ensemble models to predict BBB permeability for 1,757 chemicals using structure-derived molecular fingerprints. They input the SMILES descriptor of chemicals into the PaDEL-Descriptor software to calculate molecular fingerprints. Combining three types of ML models, including RF, SVM, and XGBoost models, this study achieved an ACC of 0.910, an ROC-AUC of 0.957, a sensitivity (SEN) of 0.927, and a specificity of 0.867 for predicting BBB permeability[75]. Recent advancements have applied more sophisticated DL methods, such as graph- and image-based models[86]. For example, Nguyen *et al.* developed a geometric multi-color message-passing graph neural network (GMC-MPNN) for BBB permeability prediction, incorporating three-dimensional geometric information and atom-type-specific subgraphs into the graph representation[76]. Their model achieved AUC-ROC values of 0.947 ± 0.011 and 0.9212 ± 0.0261 for BBB permeability classification on the MoleculeNet and B3DB benchmark datasets, respectively. For continuous permeability prediction using the B3DB dataset, the model achieved an RMSE of 0.5628 ± 0.0651 and a Pearson correlation coefficient of 0.6947 ± 0.0515, demonstrating the strong potential of graph-based DL for BBB-related pharmacokinetic (PK) characterization[76]. For BBB permeability of environmental chemicals, AI/ML models can help prioritize chemicals with potential central nervous system exposure. However, current models are mainly developed using pharmaceutical compounds, and their applicability to environmental chemicals remains unclear due to differences in chemical space and exposure-relevant chemical characteristics.

Another important biological barrier that determines chemical distribution is the placenta. AI-based predictions of placental transfer have been developed to estimate maternal-to-fetal chemical distribution for both pharmaceutical compounds and environmental chemicals. Li *et al.* developed a QSAR model to predict TTE, defined as the cord blood-to-maternal blood concentration ratio, using structural descriptors from 51 environmental chemicals, including organochlorine pesticides, PAHs, PCBs, PBDEs, and PFASs. The model identified molecular descriptors associated with placental transfer prediction and achieved moderate

Table 3. Selected AI/ML studies to predict chemical distribution parameters

Parameter (transformation; unit; endpoint type)	Input features	Chemicals	Model	Results	Ref.
General distribution					
Fraction unbound in plasma (square-root transformation; unitless; direct TK parameter)	Open-source molecular descriptors from PaDEL, OPERA, ToxPrints, and MACCS fingerprints. Recursive feature elimination with five-fold cross-validation	2,057 chemicals [training: 1,305 chemicals (650 ToxCast and 655 pharmaceuticals); external test sets: drug I (189), drug II (432), and ToxCast (97)][a]	RF regression	♦ The optimal RF model: 30 selected descriptors; Q^2 = 0.58 in five-fold cross-validation ♦ External validation: R^2 = 0.56 (drug I), 0.61 (drug II), and 0.59 (ToxCast)	Dawson *et al.*, 2021[73]
Delivery efficiency at 24 h (%ID; quantitative surrogate endpoint)	Physicochemical properties and experimental conditions	Tumor 403; heart 252; liver 341; spleen 312; lung 274; kidney 298[b]	LR, SVR, RF, XGBoost, LightGBM, DNN	♦ Best model: DNN ♦ test R^2: tumor 0.41, heart 0.42, liver 0.45, spleen 0.79, lung 0.87, kidney 0.83; test performance was similar to 5-fold CV	Mi *et al.*, 2024[74]
BBB					
BBB permeability (binary classification: Yes/No; unitless; qualitative/supporting endpoint)	Nine molecular fingerprints calculated from SMILES using PaDEL-Descriptor (EState, MACCS, PubChem, FP4, KR, AP2D, FP4C, KRC, and APC2D)	1,757 chemicals for training, 213 for external validation[b]	RF, SVM, XGBoost; ensemble model (Ensemble Top-9)	♦ Ensemble Top-9: 5-fold CV, AUC = 0.966 ± 0.011, ACC = 0.930 ± 0.013, SEN = 0.964 ± 0.013, SPE = 0.839 ± 0.037; external validation, AUC = 0.849, ACC = 0.784, SEN = 0.812, SPE = 0.712 ♦ Lower external-validation ACC suggests potential overfitting	Liu *et al.*, 2021[75]
BBB permeability: (binary classification: Yes/No; unitless Log-transformed for continuous endpoints; unitless; qualitative/supporting endpoint)	3D geometry-aware weighted colored subgraph features encoding atom-type-specific spatial interactions, combined with RDKit-derived atomic features	Three benchmark datasets: MoleculeNet classification (1,560 BBB+, 479 BBB-); B3DB classification (4,905 BBB+, 2,835 BBB-); B3DB regression (n = 1,047). All datasets were scaffold-split into training/validation/test sets at 8:1:1[b]	GMC-MPNN	♦ Classification: AUC-ROC = 0.947 ± 0.011 (MoleculeNet) and 0.9212 ± 0.0261 (B3DB). Regression: RMSE = 0.5628 ± 0.0651 and Pearson r = 0.6947 ± 0.0515 ♦ GMC-MPNN outperformed the evaluated baseline GNN models across all three datasets	Nguyen *et al.*, 2026[76]
BBB permeability: (binary classification: Yes/No; unitless; qualitative/supporting endpoint Log-transformed for continuous endpoints; unitless; quantitative surrogate endpoint)	SMILES-derived molecular embeddings from pretrained MegaMolBART; Morgan fingerprints (2,048 bits) as baseline representation	B3DB database: 7,807 compounds, including 1,058 with measured logBB CMUH-NPRL: 2,499 compounds (binary) Train: validation: test = 8:1:1[b]	MegaMolBART (pretrained LLM) molecular encoder combined with XGBoost	♦ Final combined classification model achieved AUC = 0.88 on the held-out test set ♦ MegaMolBART embeddings outperformed Morgan fingerprints for logBB regression	Huang *et al.*, 2024[77]
Placental					
TTE (unitless; direct TK parameter)	15 molecular descriptors (9 constitutional descriptors, 1 molecular property, 2 2D atom pairs, 1 2D matrix-based descriptor, 2 edge adjacency indices)	51 environmental chemicals[c]	OLS regression with stepwise	♦ Model 1 (10 selected descriptors): R_{adj}^2 = 0.67 ♦ Model 2 (5 descriptors): R_{adj}^2 = 0.61	Li *et al.*, 2021[78]
Fetal/maternal ratio (unitless; quantitative surrogate endpoint)	4 fingerprints (Morgan fingerprint, RDKit fingerprint, Hashed atom pair fingerprint, MACCS keys)	212 chemicals (environmental chemicals and 153 drugs; randomly split into training, test, and validation sets at 8:1:1)[b]	12 models: LR, decision tree, RF, SVM, KNN, XGBoost, DNN, MLP, CNN, RNN, LSTM, and transformer	♦ Retrained LSTM achieved R^2 = 0.91, 0.68, and 0.56 for the training, test, and validation sets, respectively ♦ *In vivo* evaluation of four screened chemicals showed F/M ratios > 0.3; oxybenzone had a measured F/M of 0.79 ± 0.06 *vs.* a predicted value of 0.77	Chen *et al.*, 2024[79]

Probability score to cross the placental barrier (0-0.8: low-potential; 0.8-0.9: moderate-potential; 0.9-1: high-potential; unitless; qualitative/supporting endpoint)	9 selected input features (PaDEL molecular descriptors, pathway activation score)	307 chemicals for model development; 18 compounds for external validation; 259 environmental chemicals for prospective prediction[b]	XGBoost, RF, SVM, LR, NB, and 2 ensemble models	♦ Best model: XGBoost (ACC_{test}: 0.94; F1: 0.97; AUC: 1) ♦ High-potential structures: aromatic rings and hydrophobic groups, which enhance lipid solubility ♦ Low-potential structures: nitrogen-containing carbon chains and nitrogen-containing aromatic rings, leading to lower lipid solubility ♦ The model was applied to 259 environmental chemicals to identify high-potential chemicals to transfer the placental barrier ♦ AD: evaluated using Euclidean distance-based chemical space analysis	Guan *et al.*, 2024[80]
Binding affinity to GST/NAT2 (kcal/mol; qualitative/supporting endpoint)	Molecular weight, total energy, binding energy, energy gap, ionization energy, chemical hardness, chemical softness	10 PFAS (90% training; 10% test)[c]	Multilayer perceptron-based ANN (additional molecular docking and density functional theory analyses)	♦ The ANN results showed a regression coefficient of 0.866 for GST and 0.966 for NAT2 ♦ For GST binding affinity prediction, molecular weight is the most influential factor, followed by total energy, binding energy, and other factors ♦ For NAT2 binding affinity prediction, binding energy is the most important, followed by total energy, energy gap, and other factors	Duru *et al.*, 2023[81]
Lactational					
Milk-to-plasma ratio transfer risk (binary classification; high risk: M/P ≥ 1, low risk: M/P < 1; qualitative/supporting endpoint)	5 sets of 1D & 2D molecular descriptors (MOE, DS, Mold2, RDKit, and Chemopy); fingerprints (PaDEL)	573 chemicals with experimental M/P ratios (125 high-risk, 250 low-risk); external validation: 198 chemicals detected in human milk[b]	BRF, EEC (base: AdaBoost), BBC (base: GBDT, lightGBM, XGBoost, SVM, and MLP)	♦ MOE+DS_GA_84/BRF (ACC: 81%; MCC: 61.41%; $ACC_{external}$: 86.36%) ♦ Chemopy_GA_101/BRF (ACC: 78.67%; $ACC_{external}$: 78.28%) ♦ SHAP analysis suggested the most important features of molecular hydrophobicity/hydrophilicity, π-electron, molecular polarizability, charge distribution, and conformational features	Huang *et al.*, 2025[82]
Milk-to-plasma AUC (binary classification: ≥ 1 or < 1; unitless; qualitative/supporting endpoint)	254 candidate molecular descriptors from the ADMET predictor (SMILES as the input for ADMET)	403 compounds[a]	ANN, SVM	♦ ANN: ACC_{test}: 0.929; SEN_{test}: 0.833 ♦ SVM: ACC_{test}: 0.938; SEN_{test}: 0.677 ♦ High contributions of charge-based descriptors in both models	Maeshima *et al.*, 2023[83]
Milk-to-plasma concentration ratio (binary classification: Class 1: M/P ≤ 0.1; Class 2: M/P > 0.1; unitless; qualitative/supporting endpoint)	Initially 400 descriptors for 126 drugs. Stepwise variable selection method selected the 5 most important ones [*n*-Octanol–water partition coefficient, Randic index (order 2), Max. partial charge for a C atom, Min. e–e repulsion for a C–C bond, and Min. coulombic interaction for a C–C bond]	126 drugs (96 training, 30 test; 9 external compounds without measured M/P values)[a]	SVM, LDA	♦ Best model: SVM ♦ SVM model: the classification accuracies for the training set and test set were 90.63 and 90.00%, respectively	Zhao *et al.*, 2006[84]

[a]Studies used mixed environmental and pharmaceutical datasets (mixed relevance). [b]Studies are primarily based on pharmaceutical datasets and are considered indirectly relevant to environmental TK applications. [c]Studies are based primarily on environmental chemical datasets (direct relevance). AI/ML: Artificial intelligence and machine learning; TK: toxicokinetics; OPERA: Open Structure–activity/property Relationship App; ToxPrints: Toxicological Prioritization Database fingerprints; MACCS: molecular access system; ToxCast: Toxicology in the 21st Century program; RF: random forest; %ID: percentage of injected dose; LR: linear regression; SVR: support vector regression; XGBoost: extreme gradient boosting; LightGBM: light gradient boosting machine; DNN: deep neural network; CV: cross-validation; BBB: blood–brain barrier; SMILES: simplified molecular-input line-entry system; SVM: support vector machine; AUC: area under the curve; ACC: accuracy; SEN: sensitivity; SPE: specificity;

GMC-MPNN: geometric multi-color message-passing graph neural network; ROC: receiver operating characteristic curve; RMSE: root mean squared error; GNN: graph neural network; BART: bidirectional and auto-regressive transformers; LLM: large language model; TTE: transplacental transfer efficiency; OLS: ordinary least squares; KNN: k-nearest neighbors; MLP: multilayer perceptron; CNN: convolutional neural network; RNN: recurrent neural network; LSTM: long short-term memory neural network; NB: naïve Bayes; AD: applicability domain; PFAS: per- and polyfluoroalkyl substance; ANN: artificial neural network; GST: glutathione s-transferase; BRF: balanced random forest; EEC: easy ensemble classifier; BBC: balanced bagging classifier; GBDT: gradient boosting decision tree; MCC: Matthews correlation coefficient; SHAP: Shapley Additive exPlanations; LDA: linear discriminant analysis.

predictive performance, with an adjusted R^2 of 0.67 for the 10-descriptor model and 0.61 for the 5-descriptor model[78]. More recently, Guan *et al.* incorporated placental pathway-related features (pathway activation scores derived from a placental gene network) together with molecular descriptors to classify 307 chemicals, both pharmaceutical compounds and environmental chemicals, with high or low placental barrier crossing potential, achieving high predictive performance in external validation (ACC_{test}: 0.94; F1: 0.97; AUC: 1)[80]. The study further evaluated model applicability using Euclidean distance-based chemical space analysis, showing that the prediction test set fell within the chemical space represented by the training dataset. This model showed high potential to predict the transplacental potential of environmental chemicals. However, current AI-based placental transfer models mainly predict transfer ratios or classification outcomes rather than mechanistic placental TK parameters, and the limited availability of environmentally relevant placental kinetic datasets remains a major challenge for broader application in exposure assessment.

Beyond gestation, some toxicologists have also constructed a few AI/ML models to predict relevant parameters for the chemical transmission process during lactation[87]. During this process, chemicals in maternal circulation can partition into breast milk and subsequently contribute to infant exposure. For example, Huang *et al.* developed an explainable ML model to classify chemicals with high milk-transfer risk (defined as milk-to-plasma concentration ratios ≥ 1) for 375 chemicals, with an external validation set of 198 chemicals[82]. Using molecular descriptors and fingerprints to represent physicochemical characteristics and chemical structures, respectively, the balanced RF classifier achieved the best predictive performance with an AUC of 0.87, and an ACC of 82.67% on the internal test set and 86.36% on the external test set[82]. The SHAP analysis further identified hydrophobicity/hydrophilicity, π-electron characteristics, pH-dependent lipophilicity, charge distribution, and conformational features as the key predictors of chemical milk-transfer risks. In particular, higher fractional hydrophobic surface area increased transfer possibility, whereas greater π-electron delocalization reduced it[82].

AI for metabolism

The chemical hazard may be determined not only by the parent compound itself, but also by the formation, persistence, and activity of its metabolites. In TK, metabolism occurs primarily in the liver. Depending on the chemical and exposure route, it can also occur in the intestine, kidney, lung, skin, and placenta[88]. Typically, these processes are mediated by phase I and phase II enzymes, among which cytochrome P450 enzymes, UDP-glucuronosyltransferases (UGTs), sulfotransferases (SULTs), and esterases are particularly important for the metabolic fate of xenobiotics.

AI-based prediction of metabolism-related processes requires not only characterization of enzyme–chemical interactions but also quantitative estimation of metabolic capacity that determines systemic exposure [Table 4]. Current models regarding metabolism-related TK parameters have mainly focused on endpoints such as enzyme affinity or interaction, intrinsic clearance, and metabolic stability[95]. Compared with other ADME processes, metabolism-related prediction is more strongly shaped by enzyme-specific recognition and catalytic transformation. In addition to general structural and physicochemical features, metabolism is often influenced by the accessibility of reactive sites, steric hindrance, electronic properties, and the presence of functional groups susceptible to oxidation or conjugation[96]. A representative example is the study by Sun *et al.*, who developed QSAR models to predict CYP activity profiles of environmental chemicals using models

Table 4. Selected AI/ML studies to predict chemical metabolism parameters

Parameter (transformation; unit)	Input features	Chemicals	Model	Results	Ref.
CYP activity/inhibition (binary classification; unitless; qualitative/supporting endpoint)	Atom types and correction factors	Training: 17,143 drug-like compounds from PubChem; Test: Tox21 chemicals[a]	SVM	Largely accurate for CYP1A2, CYP2C9, and CYP3A4 isozymes (AUC-ROC = 0.82-0.84) AD: evaluated using k-NN structural similarity. A large fraction of Tox21 compounds fell outside the original drug-derived AD; atom-type decomposition increased AD coverage	Sun *et al.*, 2012[89]
CYP450 SoM (atomic-site classification; unitless; qualitative/supporting endpoint)	2D topological circular fingerprints encoding SYBYL atom types at different topological distances from each candidate atom	Publicly available curated CYP3A4, CYP2D6, and CYP2C9 metabolism datasets[b]	PRW, NB, and RASCAL probabilistic classifiers	♦ PRW identified a true SoM among the top two predictions for 85%, 91%, and 88% of CYP3A4, CYP2D6, and CYP2C9 compounds, respectively; RASCAL: 83%, 91%, and 88%; NB showed lower performance (51%, 73%, and 74%) ♦ Both models had better performance than NB. PRW had a lower computational expense than RASCAL	Tyzack *et al.*, 2014[90]
CYP450 SoM for nine human CYP isoforms (atomic-site classification; unitless; qualitative/supporting endpoint)	Nine atom descriptors and four bond descriptors	Training: EBoMD (679 compounds) Test: EBoMD2 (98 compounds)[b]	D-CyPre: two D-MPNNs + XGBoost	Precision mode: Jaccard = 0.497, F1 = 0.660, precision = 0.737; recall mode: Jaccard = 0.506, F1 = 0.669, recall = 0.720. D-CyPre outperformed BioTransformer for all nine CYP isoforms and CyProduct for 5/9 isoforms	Yang *et al.*, 2024[91]
Human liver microsomal metabolic stability (binary classification: stable ≥ 50% remaining at 30 min; unstable < 50%; unitless; qualitative/supporting endpoint)	1,249 mordred molecular descriptors calculated from SMILES; RF selected 200 descriptors	1,917 experimentally measured compounds (1,049 stable, 868 unstable); Train: test = 8:2; additional experimentally measured external test set N = 61[b]	ANN, kNN, LR, NB, RF, SVM	♦ Best model: RF ♦ Five-fold CV AUC-ROC = 0.73 ♦ Internal test: ACC = 0.68, MCC = 0.34, SEN = 0.76, SPE = 0.57 ♦ External test: ACC = 0.74, MCC = 0.48, SEN = 0.70, SPE = 0.86, PPV = 0.94, NPV = 0.46 ♦ Chemical-space/generalizability: external compounds were intentionally structurally distinct from training compounds; maximum Tanimoto similarity was only 0.53	Ryu *et al.*, 2022[92]
Metabolic stability & CYP inhibition category (classification; unitless; qualitative/supporting endpoint)	Molecular structural descriptors combined with features related to prediction confidence and structural similarity	26,138 compounds for metabolic stability 16,613 compounds for CYP inhibition[b]	Hybrid ML framework with a separate AD classifier	♦ Models trained on public datasets transferred poorly to structurally different in-house compounds ♦ AD: structural similarity and prediction probability; a separate ML model classified new compounds as inside or outside the AD	Sasahara *et al.*, 2021[93]
Hepatic intrinsic clearance (log-transformed; direct TK parameter)	17-65 descriptors derived from 1,710 structural and physicochemical descriptors (RDKit and Mordred)	212 chemicals[a]	LightGBM	♦ r = 0.77 (P < 0.01, AAFE = 3.28) ♦ Applying the predicted values of $CL_{h,int}$ and two other parameters (absorption rate constants and volumes of the systemic circulation) to a simplified PBPK model, the r for logC_{max} and logAUC were 0.85 and 0.80, respectively	Kamiya *et al.*, 2022[94]
Intrinsic clearance, CL_{int} (categorical; μL/min/10^6 cells; qualitative/supporting endpoint)	PaDEL, OPERA, ToxPrints, and MACCS descriptors/fingerprints	3,713 chemicals from ToxCast and ChEMBL; N_{train} = 1,600 with independent ToxCast and ChEMBL test sets[a]	RF classification	Best 3-bin model: test ACC = 0.704 for ToxCast and 0.469 for ChEMBL after AD filtering	Dawson *et al.*, 2021[73]

[a]Studies used mixed environmental and pharmaceutical datasets (mixed relevance). [b]Studies are primarily based on pharmaceutical datasets and are considered indirectly relevant to environmental TK applications. AI/ML: Artificial intelligence and machine learning; CYP: cytochrome P450; SVM: support vector machine; AUC-ROC: area under the receiver operating characteristic curve; AD: applicability domain; k-NN: k-nearest neighbors; SoM: site of metabolism; PRW: Parzen–Rosenblatt Window; NB: naïve Bayes; RASCAL: Random Attribute Subsampling Classification Algorithm; D-MPNN: directed message-passing neural network; XGBoost: extreme gradient boosting; SMILES: simplified molecular-input line-entry system; RF: random forest; ANN: artificial neural network; LR: logistic regression; ACC: accuracy; MCC: Matthews correlation coefficient; SEN: sensitivity; SPE: specificity; PPV: positive predictive value; NPV: negative predictive value; TK: toxicokinetics; LightGBM: light gradient boosting

machine; AAFE: average absolute fold error; PBPK: physiologically based pharmacokinetic; Cmax: maximum concentration; OPERA: Open Structure–activity/property Relationship App; MACCS: molecular access system; ChEMBL: Chemical Database of the European Molecular Biology Laboratory.

built on drug-like compounds and then applied them to the environmental chemical space[89]. Using the activity/inhibition data across different CYP isoforms for 17,143 drug-like compounds from PubChem, the researchers developed an SVM model to predict the CYP isozymes and applied it to Tox21 compounds. The predictions were largely accurate for CYP1A2, CYP2C9, and CYP3A4 isozymes with AUC-ROC ranging between 0.82 and 0.84[89].

AI methods have also been applied to more mechanistic metabolism endpoints, such as site-of-metabolism and metabolite prediction. For example, Tyzack *et al.* developed probabilistic classifiers for CYP450 site-of-metabolism prediction using 2D topological fingerprints, showing that machine-learning approaches can identify likely metabolic sites at the atomic level[96]. More recently, deep-learning models have been used to predict likely metabolite structures and reaction outcomes, extending metabolism modeling beyond simple clearance-related endpoints toward explicit biotransformation prediction[97]. Yang *et al.* developed D-CyPre, a deep-learning framework, to predict the metabolic sites across nine human CYP450 isoforms[91]. This model used directed message-passing neural networks to integrate atom-, bond-, and molecular-level structural information, achieving F1 scores of 0.66-0.67 on the test set and outperformed several existing metabolism predictors[91]. These approaches provide mechanistic insights into potential metabolic pathways and may help identify metabolites with different persistence or toxicity profiles. Nevertheless, prediction of metabolite formation remains challenging because of the diversity of enzyme systems, multiple competing metabolic pathways, and limited availability of experimentally characterized metabolite datasets.

Beyond enzyme-level activity and metabolite formation, AI/ML approaches have also been applied to predict quantitative metabolism-related TK parameters, particularly hepatic intrinsic clearance ($CL_{int,h}$), which is a key determinant of hepatic metabolism and systemic exposure in TK models. This parameter is particularly important because it reflects the intrinsic metabolic capacity of the liver independent of hepatic blood flow and plasma protein binding[98]. It is primarily determined by the activity and abundance of hepatic drug-metabolizing enzymes and therefore represents the rate at which the liver can metabolically eliminate a chemical when delivery to the liver is not limiting. Under nonsaturating conditions, $CL_{int,h}$ is commonly combined with hepatic blood flow and the fraction unbound in blood or plasma to estimate hepatic clearance in mechanistic TK and PBPK modeling. Because $CL_{int,h}$ is chemical-specific but often unavailable experimentally for many environmental chemicals, it has become an important target for *in silico* and AI-based prediction. For example, Kamiya *et al.* predicted the hepatic intrinsic clearance of 212 chemicals (both environmental xenobiotics and medicines), along with absorption rate constants and the systemic circulation volumes, using 17-65 structural and physicochemical descriptors selected from 1,710 calculated descriptors[94]. The LightGBM model achieved a correlation coefficient of 0.77 and an average absolute fold error of 3.28 for hepatic intrinsic clearance. The three predicted parameters were subsequently incorporated into a simplified PBPK model. The maximum plasma concentrations and areas under the concentration–time curve generated using the *in silico*-estimated parameters were correlated with those generated using traditionally determined parameters, with correlation coefficients of 0.85 and 0.80, respectively[94]. Rather than treating model outputs as standalone predictions, this study's workflow showed that the predicted parameters can be incorporated into a mechanistic TK or PBPK framework, where their combined impact can be evaluated against observed concentration–time data of environmental chemicals.

AI for excretion

Chemical excretion is the irreversible removal of xenobiotics or their metabolites from the body and is a key determinant of overall TK behavior[99]. Environmental chemicals are excreted through renal and biliary pathways in the general population. Other routes may include exhalation, saliva, mucosal, and sweat,

depending on the compound's physicochemical properties[16].

The primary excretion pathway depends on the chemical molecular structure and physicochemical features. For polar and ionizable environmental chemicals, renal excretion is typically the dominant pathway, which involves glomerular filtration, active tubular secretion, and reabsorption processes[100]. Besides the physiological features common to the other three processes, several specific molecular determinants may influence renal clearance, such as interaction efficiencies with organic cation transporter 2 (OCT2), organic anion transporters (OAT1/3), and multidrug and toxin extrusion proteins (MATEs)[100-103]. For example, the chemicals within the PFAS family exhibit substantial variability in renal elimination rates, driven by differences in carbon chain length, protein-binding fractions, and interactions with transporters[104].

In contrast, some chemicals are primarily eliminated through the biliary pathway, particularly those with relatively large molecular weight (typically > 300-500 Da) and higher lipophilicity (e.g., logP > 2-3). A statistically derived threshold suggests that compounds with molecular weight above approximately 350 Da are more likely to undergo significant biliary elimination[105]. This process requires active secretion of chemicals and/or their metabolites from hepatocytes into bile, followed by elimination in feces, which is mediated by hepatic transporters and influenced by molecular weight, polarity, and conjugation status[106-108].

These mechanistic determinants form the basis for AI-driven prediction of excretion-related parameters. Renal excretion has received comparatively more attention, with models developed for quantitative renal clearance (CL_R) and the fraction of a chemical excreted unchanged in urine (fe). For example, Paine *et al.* developed 3 statistical models [partial least squares, RF, as well as classification and regression trees (CARTs)] to predict the human renal clearance rate of 349 compounds [acids (N = 59), bases (N = 124), zwitterions (N = 112), and neutral (N = 54)][109]. Using the molecular structures of these compounds, a total of 195 molecular descriptors were generated to characterize the key molecular features, including lipophilicity, hydrogen-bonding, size, charge/polarity, and topology. Among the tested models, the RF model showed moderate to high predictive performance across compound classes, with particularly strong performance for acids and zwitterions (r_{obf}^2 = 0.79). Top influencing descriptors included lipophilicity-based and positive charge-related descriptors[109]. A later study developed a two-stage *in silico* framework for human renal excretion using datasets of 411 compounds for fe and 401 compounds for CL_R[110]. The first model classified whether a compound was predominantly renally excreted, achieving a balanced ACC of 0.74. The subsequent models further separated compounds by renal excretion type and incorporated the fraction unbound in plasma ($f_{u,p}$) as an additional descriptor. Inclusion of measured or predicted $f_{u,p}$ improved quantitative clearance prediction, with 78.6% of compounds in the high-CL_R group predicted within twofold error when predicted $f_{u,p}$ was used[110]. This example illustrates the advantage of incorporating a physiologically relevant TK parameter into structure-based renal clearance prediction rather than relying on molecular structure alone.

Compared with renal clearance, AI-based prediction of biliary excretion remains less developed, partly because quantitative human biliary-excretion data are relatively limited and transporter-mediated processes are difficult to characterize experimentally[111]. One representative QSAR study predicted the percentage of biliary excretion of 217 compounds using molecular descriptors. A simple regression tree model generated using the CART algorithm provided the best predictive performance (MAE_{test} = 0.373), and model interpretation indicated that larger molecular size, ionic character, and moderate lipophilicity favored biliary excretion. The study also identified a statistically supported molecular-weight threshold of approximately 348 Da for significant biliary excretion and found poorer model performance for compounds with extreme lipophilicity (logP > 5.35) or molecular weight below approximately 280 Da.

Overall, current AI applications for excretion appear more mature for renal endpoints than for biliary elimination. Renal models benefit from relatively well-defined quantitative endpoints such as renal clearance and fraction excreted unchanged in urine, whereas biliary excretion involves sequential hepatic uptake, intracellular handling, transporter-mediated canalicular secretion, and possible enterohepatic recirculation, making a single structure–endpoint relationship more difficult to establish. In addition, most existing models were developed primarily from pharmaceutical datasets, and their generalizability to environmental chemical space remains insufficiently evaluated. Future models may therefore benefit from integrating molecular descriptors with transporter-specific information and other mechanistically relevant TK parameters rather than relying on chemical structure alone.

PERSPECTIVE AND FUTURE DIRECTIONS

As described above, AI/ML approaches are increasingly used to improve TK parameterization and expand prediction coverage for chemicals lacking experimental data. By enabling more efficient description and estimation of ADME processes, these approaches help bridge the gap between external exposure and internal dose, thereby not only speeding up TK parameterization but also facilitating the understanding of the toxicological impacts. However, as the field moves from parameter prediction toward practical application in exposure reconstruction and risk assessment, several methodological and application-related challenges must be addressed to ensure that AI/ML-derived TK predictions are reliable and practically useful.

A key methodological consideration is whether AI/ML models can provide reliable predictions beyond the datasets on which they were developed. Similar statistical performance can represent very different levels of practical confidence depending on the validation strategy, chemical-space coverage and AD, uncertainty characterization, and, where relevant, downstream TK/PBPK evaluation. We therefore compared these features, together with environmental relevance, across representative studies in Table 5.

Among these considerations, overfitting remains an important source of reduced model reliability and generalizability. High predictive performance obtained from random train–test splits or conventional cross-validation does not necessarily indicate that a model will perform well for structurally distinct chemicals, particularly when closely related analogs occur in both training and test sets. Therefore, assessment of AI/ML-based TK models should extend beyond conventional performance metrics and consider validation design, chemical-space overlap, AD, and prediction uncertainty. These considerations are consistent with OECD guidance for QSAR model validation, which emphasizes the need for a defined AD and appropriate measures of model robustness and predictivity[61]. Several studies reviewed here illustrate different approaches to this issue. For oral bioavailability prediction, Ng and Lu used repeated five-fold cross-validation and early stopping to reduce overfitting during GNN training. However, the model's robustness for compounds outside the training chemical space remained insufficiently evaluated, indicating that controlling overfitting during training does not necessarily ensure reliable extrapolation to structurally distinct chemicals[44]. Beyond internal strategies for controlling overfitting, some studies used more stringent data-splitting approaches to evaluate model generalizability. Obrezanova *et al.* used a temporal test split to predict rat *in vivo* PK parameters and further examined scaffold overlap between the training and test sets[113]. Dawson *et al.* explicitly defined the AD of their QSAR models using standardized ranges of training-set descriptors and removed out-of-domain test chemicals before calculating independent test-set performance[73]. Similarly, Sun *et al.* evaluated the transfer of CYP models developed primarily from drug-like compounds to Tox21 environmental chemicals and showed that a substantial proportion of Tox21 compounds fell outside the original model AD, illustrating that acceptable overall predictive performance can coexist with incomplete chemical-space coverage[89]. Together, these examples indicate that controlling overfitting during model development and evaluating model generalizability after development are related but distinct requirements. Internal approaches such as cross-validation, regularization, and early stopping

Table 5. Evaluation of model reliability in representative AI/ML applications for environmental TK

TK parameter (ADME)	Database environmental relevance	External validation	AD/chemical-space assessment	Uncertainty assessment	Mechanistic integration	Key consideration for model reliability
Oral bioavailability (Absorption)[44]	Indirect - Primarily drug-like compounds	Repeated 5-fold CV, independent test set	t-SNE analysis train/test within same chemical space; No OOD test	NR	No	Transfer learning improved predictive performance, but generalizability to structurally distinct environmental chemicals remains insufficiently evaluated
Fraction unbound in plasma (Distribution)/intrinsic clearance (Metabolism)[73]	Mixed - pharmaceutical and ToxCast chemicals	Independent pharmaceutical and ToxCast test sets	Standardized descriptor ranges	Assay/model/exposure uncertainty discussed	Httk reverse dosimetry/BER	Independent testing together with explicit AD assessment provides stronger evidence for application to data-poor environmental chemicals
BBB permeability (Distribution)[76]	Indirect - drug-oriented benchmark dataset	5 scaffold-based 8:1:1 splits	NR	Mean ± SD across 5 splits	No	Strong structural generalization test
Placental barrier crossing (Distribution)[80]	Mixed/direct application - mixed model-development set followed by prediction of environmental chemicals	Independent external validation set (n = 18)	Euclidean distance	NR	No	Combines external validation with chemical-space analysis and prospective application to environmental chemicals
CYP activity/inhibition (Metabolism)[89]	Mixed/ direct application - drug-like training data applied to Tox21 chemicals	Tox21 chemicals used to evaluate model transfer	Assessment using structural similarity	NR	No	Acceptable predictive performance; incomplete AD coverage
Hepatic intrinsic clearance (Metabolism)[94]	Mixed - environmental chemicals and medicines	No independent environmental external set reported	NR	Prediction error reported (e.g., AAFE)	Yes - a simplified PBPK model integration	Downstream PBPK evaluation provides additional evidence that parameter-level predictions can support concentration–time simulation
Human renal clearance (Excretion)[109]	Indirect - primarily pharmaceutical compounds	Hold-out testing across chemical classes	PCA; train/test property-space overlap	Y-permutation tests	No	Robust internal validation; OOD generalization untested
Dermal partition and diffusion parameters (Absorption)[68]	Direct - 43 environmental chemicals	80:20 train/test + 11-chemical experimental validation	Limited chemical space; no full external AD assessment	Mean ± SD; factor-of-2 comparison	Yes - mechanistic diffusion model	Predicted parameters were evaluated through downstream dermal permeation simulations, but broader external validation is still needed
NP tumor kinetic parameters (Distribution)[112]	Indirect - Drug-loading NPs	Nested 5-fold CV + independent test set	NR	CV variability + prediction error	Yes - PBPK integration	QSAR and PBPK components were developed from the same dataset with consistent parameter definitions, allowing direct integration of predicted parameters into the PBPK model

Nanoparticle tissue concentration at 24 h (Distribution)[74]	Indirect - Drug-loading NPs	80:20 independent test + 5-fold CV	NR	CV variability + prediction error	No	Consistent CV/test performance; multi-tissue validation

Environmental relevance was categorized as direct when the model-development or application dataset predominantly involved environmental chemicals or emerging environmental contaminants, mixed when both environmental and pharmaceutical chemicals were included, and indirect when models were developed primarily from pharmaceutical datasets. AI/ML: Artificial intelligence and machine learning; TK: toxicokinetics; ADME: absorption, distribution, metabolism, and excretion; CV: cross-validation; t-SNE: t-distributed stochastic neighbor embedding; OOD: out of domain; NR: not reported; BER: bioactivity–exposure ratio; AD: applicability domain; BBB: blood–brain barrier; SD: standard deviation; AAFE: average absolute fold error; PBPK: physiologically based pharmacokinetic; NP: nanoparticle; QSAR: quantitative structure-activity relationship.

can reduce overfitting, whereas scaffold- or cluster-based splitting, temporal validation, independent external testing, and explicit AD analysis provide more informative evidence of whether a model can be applied beyond its training chemical space.

Beyond model reliability and generalizability, another important consideration is interpretability and explainability of AI/ML models. Interpretability should be viewed as a prerequisite for future environmental TK modeling because the "black box" nature of AI/ML models can limit the understanding of how predictions are generated and decisions are made, and subsequently hinder regulatory and scientific acceptance. To address this limitation, a key future need for AI-based TK parameterization in environmental exposure assessment is to improve model explainability. Because different scientific communities have different prediction tasks, no universal criteria exist for interpretable ML[114]. In the context of TK parameterization and exposure assessment, we consider that model explainability can be advanced through two complementary forms of interpretability proposed by Tjoa and Guan: perceptive interpretability and interpretability by mathematical structure[115,116]. Perceptive interpretability can be achieved through saliency-based approaches that explain model decisions by assigning values reflecting the relative importance of input features[115]. In AI/ML models for TK parameterization, these approaches help identify which input structural and physicochemical features drive the prediction. For example, in the human milk transfer model discussed above, explainability analysis (SHAP analysis) suggested that hydrophobicity/ hydrophilicity and pH-dependent lipophilicity were among the major features influencing the predictions of milk transfer risks[82]. However, SHAP and similar feature-attribution methods explain how inputs influence model predictions but do not establish biological causality or provide direct evidence of underlying mechanisms. In contrast, interpretability can be improved through mathematical structure by using models that explicitly describe relationships between input features and predicted outcomes or incorporate known biological processes, such as additive response functions, symbolic equations, or mechanistically structured hybrid models[115]. This type of interpretability is particularly valuable for TK parameterization because it can further quantify how a given input descriptor, such as a structural or physicochemical feature, relates to predicted ADME property values and subsequent internal exposure in humans or specific populations. More broadly, several related strategies have been developed and applied in TK- and PK-oriented AI/ML modeling to improve these two forms of model interpretability, such as feature attribution, functional extraction, hybrid mechanistic-ML models, and uncertainty-aware modeling with domain adaptation[45,114]. Nevertheless, model explainability alone is insufficient for regulatory acceptance, which also requires external validation, a clearly defined AD, uncertainty characterization, and a specified context of use.

Another important future consideration is the continued expansion of data resources, specifically for TK modeling of environmental chemicals. As described above and shown in Tables 2 and 3, more than half of the studies developed AI/ML models using PK data, partly because both public and in-house data resources are much more abundant in the pharmaceutical setting than those in environmental toxicology. TK data for environmental chemicals remain relatively limited, fragmented, and often less standardized. Except for a few government databases, researchers would struggle to build large-scale TK databases, which would limit

AI/ML model development and reduce the model's predictive capacity. Although TK and PK share many underlying ADME principles, TK parameterization studies focus more on environmental exposure scenarios, along with the chemical itself. As a result, PK-based databases cannot always capture the chemical space, dose ranges, biological matrices, and toxicity-relevant endpoints that are more informative for environmental TK assessment. Future progress in AI-based TK parameterization will benefit not only from larger databases, but also from more TK-specific curation strategies that better represent environmental chemicals and exposure-relevant contexts.

Finally, integrating AI/ML prediction results, such as AI-derived TK parameters, into downstream mathematical modeling will support environmental assessment by translating data-driven predictions into mechanistically interpretable exposure assessments. For example, AI-predicted TK parameters, such as clearance, fraction unbound, permeability-related metrics, and partition coefficients, can be incorporated into PBPK models to simulate the concentration-time changes of environmental chemicals in different tissues. In such applications, AI models primarily serve as tools to estimate the key parameters, while the integrated PBPK model can provide a mechanistic framework to integrate these parameters with physiological processes and predict chemical concentrations in target tissues. When AI models are trained using *in vivo*-derived TK parameters, the predicted values may be incorporated directly into PBPK models after evaluating data quality and confirming compatibility with the corresponding PBPK parameter definitions and model requirements. In contrast, when AI models predict parameters from *in vitro* datasets, these parameters generally require *in vitro*-to-*in vivo* extrapolation or other biological scaling approaches before being used as PBPK inputs. Therefore, the reliability of AI–PBPK modeling depends not only on the ACC of individual parameter predictions but also on appropriate parameter translation, uncertainty characterization, physiological consistency, and validation against observed kinetic profiles. A few studies have developed AI/ML-based PBPK models for PK profile simulation[45,65,94,117-119]. For example, Chou *et al.* developed an AI-based PBPK model in which a QSAR model predicted four critical nanoparticle kinetic parameters in the tumor microenvironment, and these predicted values were incorporated into a pre-validated PBPK model[112]. In this study, the QSAR and PBPK components were developed from the same underlying kinetic dataset, ensuring consistency in parameter definitions and modeling context and thereby allowing the QSAR-predicted parameters to be directly used as PBPK inputs. The PBPK model simulation of drug concentration in the target tissue (i.e., the tumor) was used for the AI model evaluation because these four parameters were not readily measured[45]. More recently, Mi *et al.* applied an AI-assisted PBPK model to derive nanoparticle tumor-delivery kinetics under administration regimens used in pharmacodynamic experiments. ML-predicted tumor-related PBPK parameters and administration regimens were incorporated into the PBPK model to simulate tumor concentration–time profiles and derive PK metrics, including AUC_{tumor}, DE_{24}, and DE_{max} [30]. These PK metrics were then combined with nanoparticle physicochemical and experimental features in ML models for antitumor efficacy prediction, illustrating an integrated AI–PBPK–PD workflow. In the future, when AI-predicted TK parameters are integrated into human PBPK models to simulate chemical exposure over time, population variability should be further considered by incorporating inter-individual differences in physiological characteristics (e.g., body weight, organ volumes, and blood flows) and toxicokinetic processes (e.g., metabolic capacity, renal clearance, and protein binding).

CONCLUSION

In summary, AI/ML modeling is increasingly being used for TK parameterization and environmental exposure assessment. By enabling scalable prediction of ADME-related parameters, AI/ML approaches can help address the major data gap caused by the large number of environmental chemicals and the limited throughput of conventional *in vivo* and *in vitro* assays. As summarized in this review, recent progress has expanded from classical descriptor-based models to advanced DL architectures and has been applied across the major ADME processes, with increasing applicability to environmental chemicals across exposure routes.

These developments are particularly valuable in environmental exposure science, where the number and diversity of chemicals greatly exceed the capacity of conventional experiments for TK characterization.

Nevertheless, the broader impact of AI-based TK parameterization on environmental exposure assessment will depend on whether these models can move beyond isolated prediction tasks and support mechanistic insights. Currently, the maturity of AI-derived TK parameters for practical application varies substantially across different endpoints. Predictions for relatively well-established parameters, such as clearance, fraction unbound, permeability-related metrics, and partition coefficients, are closer to practical application because of greater data availability and clearer links to established mechanistic models. However, AI applications for other less-characterized parameters and more complex TK processes involving multiple interacting biological mechanisms remain largely exploratory due to limited datasets, heterogeneous experimental conditions, and challenges in model validation. Future progress will require continued improvement in data quality and curation, better representation of environmental chemicals and exposure-relevant covariates, stronger model interpretability and uncertainty quantification, and more effective integration of AI-predicted parameters into PBPK and related downstream frameworks. With these advances, AI/ML has the potential not only to accelerate TK parameter estimation, but also to strengthen the linkage between external exposure, internal dose, and toxicological outcomes, thereby contributing to a more efficient and scientifically grounded paradigm for environmental risk assessment.

DECLARATIONS

Acknowledgments

We thank BioRender.com for providing the platform used to create the graphical abstract [Created in BioRender. Zhang, Z. (2026) https://BioRender.com/ei8i3j5].

Authors' contributions

Conceptualization, outline development, writing - original draft: Wu, X.
Writing the original draft: Zhang, Z.
Draft revision: Mi, K.
Conceptualization, outline development, writing - original draft, writing - review and editing: Chen, Q.

Availability of data and materials

Not applicable.

AI and AI-assisted tools statement

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

Chen, Q. is the Assistant Guest Editor of the Special Issue "AI in Environmental Exposure and Health" of the journal *Journal of Environmental Exposure Assessment*. Chen, Q. was not involved in any steps of editorial processing, notably including reviewers' selection, manuscript handling and decision-making. The other authors declare no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

REFERENCES

1. National Research Council. *Toxicity testing in the 21st century: a vision and a strategy*. National Academies Press; 2007. DOI
2. Chen, Q.; Chou, W. C.; Lin, Z. Integration of toxicogenomics and physiologically based pharmacokinetic modeling in human health risk assessment of perfluorooctane sulfonate. *Environ. Sci. Technol.* **2022**, *56*, 3623-33. DOI PubMed
3. Wambaugh, J. F.; Hughes, M. F.; Ring, C. L.; et al. Evaluating in vitro-in vivo extrapolation of toxicokinetics. *Toxicol. Sci.* **2018**, *163*, 152-69. DOI PubMed PMC
4. Hartung, T. Toxicology for the twenty-first century. *Nature* **2009**, *460*, 208-12. DOI PubMed
5. U.S. EPA. TSCA chemical substance inventory. https://www.epa.gov/tsca-inventory. (accessed 2026-09-21).
6. Wambaugh, J. F.; Wetmore, B. A.; Pearce, R.; et al. Toxicokinetic triage for environmental chemicals. *Toxicol. Sci.* **2015**, *147*, 55-67. DOI PubMed PMC
7. Breen, M.; Ring, C. L.; Kreutz, A.; Goldsmith, M. R.; Wambaugh, J. F. High-throughput PBTK models for in vitro to in vivo extrapolation. *Expert. Opin. Drug. Metab. Toxicol.* **2021**, *17*, 903-21. DOI PubMed PMC
8. Isaacs, K. K.; Egeghy, P.; Dionisio, K. L.; et al. The chemical landscape of high-throughput new approach methodologies for exposure. *J. Expo. Sci. Environ. Epidemiol.* **2022**, *32*, 820-32. DOI PubMed PMC
9. Vamathevan, J.; Clark, D.; Czodrowski, P.; et al. Applications of machine learning in drug discovery and development. *Nat. Rev. Drug. Discov.* **2019**, *18*, 463-77. DOI PubMed PMC
10. Koirala, M.; Yan, L.; Mohamed, Z.; DiPaola, M. AI-integrated QSAR modeling for enhanced drug discovery: from classical approaches to deep learning and structural insight. *Int. J. Mol. Sci.* **2025**, *26*, 9384. DOI PubMed PMC
11. Chou, W. C.; Li, M.; Lin, Z. Chapter 4 - Application of machine learning and artificial intelligence methods in physiologically based pharmacokinetic modeling. In *Machine learning and artificial intelligence in toxicology and environmental health*. Elsevier; 2026. pp. 99-138. DOI
12. Ajisafe, O. M.; Adekunle, Y. A.; Egbon, E.; Ogbonna, C. E.; Olawade, D. B. The role of machine learning in predictive toxicology: a review of current trends and future perspectives. *Life. Sci.* **2025**, *378*, 123821. DOI PubMed
13. Lones, M. A. Avoiding common machine learning pitfalls. *Patterns* **2024**, *5*, 101046. DOI PubMed PMC
14. Wang, H.; Chen, J.; Liu, W.; et al. Using machine learning for green substitution of industrial chemicals: integrating functionality, hazard, and life cycle impact. *Chem. Rev.* **2026**, *126*, 841-94. DOI PubMed
15. Fu, X.; Wojak, A.; Neagu, D.; Ridley, M.; Travis, K. Data governance in predictive toxicology: a review. *J. Cheminform.* **2011**, *3*, 24. DOI PubMed PMC
16. Deepika, D.; Kumar, V. The role of "physiologically based pharmacokinetic model (pbpk)" new approach methodology (nam) in pharmaceuticals and environmental chemical risk assessment. *Int. J. Environ. Res. Public. Health.* **2023**, *20*, 3473. DOI PubMed PMC
17. Claire, T.; Sean, H. Integrating toxicokinetics into toxicology studies and the human health risk assessment process for chemicals: Reduced uncertainty, better health protection. *Regul. Toxicol. Pharmacol.* **2022**, *128*, 105092. DOI PubMed
18. Kim, S. J.; Heo, S. H.; Lee, D. S.; Hwang, I. G.; Lee, Y. B.; Cho, H. Y. Gender differences in pharmacokinetics and tissue distribution of 3 perfluoroalkyl and polyfluoroalkyl substances in rats. *Food. Chem. Toxicol.* **2016**, *97*, 243-55. DOI PubMed
19. Karmaus, A. L.; Kreutz, A. L.; Oyetade, O.; et al. Perspectives on variability of in vivo toxicology studies: considerations for next-generation toxicology. *Front. Toxicol.* **2026**, *8*, 1778353. DOI PubMed PMC
20. Napoli, J. A.; Reutlinger, M.; Brandl, P.; Wang, W.; Hert, J.; Desai, P. Multitask deep learning models of combined industrial absorption, distribution, metabolism, and excretion datasets to improve generalization. *Mol. Pharm.* **2025**, *22*, 1892-900. DOI PubMed
21. Kreutz, A.; Chang, X.; Hogberg, H. T.; Wetmore, B. A. Advancing understanding of human variability through toxicokinetic modeling, in vitro-in vivo extrapolation, and new approach methodologies. *Hum. Genomics.* **2024**, *18*, 129. DOI PubMed PMC
22. Alves, V. M.; Auerbach, S. S.; Kleinstreuer, N.; et al. Curated data in - trustworthy *in silico* models out: the impact of data quality on the reliability of artificial intelligence models as alternatives to animal testing. *Altern. Lab. Anim.* **2021**, *49*, 73-82. DOI PubMed PMC
23. Pearce, R. G.; Setzer, R. W.; Strope, C. L.; Wambaugh, J. F.; Sipes, N. S. httk: R package for high-throughput toxicokinetics. *J. Stat. Softw.* **2017**, *79*, 1-26. DOI PubMed PMC
24. Wambaugh, J. F.; Wetmore, B. A.; Ring, C. L.; et al. Assessing toxicokinetic uncertainty and variability in risk prioritization. *Toxicol. Sci.* **2019**, *172*, 235-51. DOI PubMed PMC
25. Mansouri, K.; Martin, T.; Chang, X.; Williams, A. J.; Allen, D.; Kleinstreuer, N. 4.18.T-01 - OPERA: open-source QSAR models for regulatory support. 2023. https://studio.m-anage.com/setac/sna2023/meetingapp.cgi/Paper/17094. (accessed 2026-09-21).
26. Chakraborty, S.; Boyina, H. K.; Mitta, R.; Nayaka, R. , In silico toxicokinetics. In *Computer simulations in the pharmaceutical industry*. CRC Press; 2026. pp. 199-220. DOI
27. Chen, C. Y.; Lin, Z. Exploring the potential and challenges of developing physiologically-based toxicokinetic models to support human health risk assessment of microplastic and nanoplastic particles. *Environ. Int.* **2024**, *186*, 108617. DOI PubMed
28. Wilhelm, S.; Tavares, A. J.; Dai, Q.; et al. Analysis of nanoparticle delivery to tumours. *Nat. Rev. Mater.* **2016**, *1*, 16014. DOI

29. Chen, Q.; Yuan, L.; Chou, W. C.; et al. Meta-analysis of nanoparticle distribution in tumors and major organs in tumor-bearing mice. *ACS. Nano.* **2023**, *17*, 19810-31. DOI PubMed PMC
30. Mi, K.; Chen, Q.; Yuan, L.; et al. Analysis of pharmacokinetic-pharmacodynamic relationships of nanoparticles against tumors. *ACS. Nano.* **2026**, *20*, 22485-504. DOI PubMed PMC
31. Alexandropoulos, S. A. N.; Kotsiantis, S. B.; Vrahatis, M. N. Data preprocessing in predictive data mining. *Knowl. Eng. Rev.* **2019**, *34*, e1. DOI
32. Cabello-Solorzano, K.; Ortigosa de Araujo, I.; Peña, M.; Correia, L.; Tallón-Ballesteros, A. J. The impact of data normalization on the accuracy of machine learning algorithms: a comparative analysis. In 18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023). Springer, Cham; 2023. pp. 344-53. DOI
33. David, L.; Thakkar, A.; Mercado, R.; Engkvist, O. Molecular representations in AI-driven drug discovery: a review and practical guide. *J. Cheminform.* **2020**, *12*, 56. DOI PubMed PMC
34. Deng, J.; Yang, Z.; Wang, H.; Ojima, I.; Samaras, D.; Wang, F. A systematic study of key elements underlying molecular property prediction. *Nat. Commun.* **2023**, *14*, 6395. DOI PubMed PMC
35. Handa, K.; Hirano, M.; Kageyama, M.; Bender, A. Computational approaches to DMPK: a realistic assessment of current methods and their practical impact. Part I: Physicochemical and in vitro properties. *Drug. Discov. Today.* **2025**, *30*, 104422. DOI PubMed
36. Ryu, J. Y.; Jang, W. D.; Jang, J.; Oh, K. S. PredAOT: a computational framework for prediction of acute oral toxicity based on multiple random forest models. *BMC. Bioinformatics.* **2023**, *24*, 66. DOI PubMed PMC
37. Xu, Y.; Pei, J.; Lai, L. Deep learning based regression and multiclass models for acute oral toxicity prediction with automatic chemical feature extraction. *J. Chem. Inf. Model.* **2017**, *57*, 2672-85. DOI PubMed
38. Jiang, J.; Wang, R.; Wei, G. W. GGL-Tox: geometric graph learning for toxicity prediction. *J. Chem. Inf. Model.* **2021**, *61*, 1691-700. DOI PubMed PMC
39. Noga, M.; Jurowski, K. Preliminary prediction of toxicologically relevant physicochemical properties of Novichoks: the first comparative in silico studies. *Chem. Biol. Interact.* **2025**, *419*, 111644. DOI PubMed
40. Li, T.; Chen, X.; Tong, W. Bridging organ transcriptomics for advancing multiple organ toxicity assessment with a generative AI approach. *NPJ. Digit. Med.* **2024**, *7*, 310. DOI PubMed PMC
41. Führer, F.; Gruber, A.; Diedam, H.; Göller, A. H.; Menz, S.; Schneckener, S. A deep neural network: mechanistic hybrid model to predict pharmacokinetics in rat. *J. Comput. Aided. Mol. Des.* **2024**, *38*, 7. DOI PubMed
42. Wu, X.; Wu, P. Y.; Chou, W. C.; Tell, L. A.; Lin, Z. A machine learning-empowered quantitative structure-activity relationship model for predicting the plasma half-life of drugs in dogs. *AAPS. J.* **2025**, *28*, 22. DOI PubMed
43. Fan, N.; Chen, J.; Wang, J.; Chen, Z. S.; Yang, Y. Bridging data and drug development: machine learning approaches for next-generation ADMET prediction. *Drug. Discov. Today.* **2025**, *30*, 104487. DOI PubMed
44. Ng, S. S. S.; Lu, Y. Evaluating the use of graph neural networks and transfer learning for oral bioavailability prediction. *J. Chem. Inf. Model.* **2023**, *63*, 5035-44. DOI PubMed PMC
45. Chou, W. C.; Lin, Z. Machine learning and artificial intelligence in physiologically based pharmacokinetic modeling. *Toxicol. Sci.* **2023**, *191*, 1-14. DOI PubMed PMC
46. Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J. C.; Sheridan, R. P.; Feuston, B. P. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947-58. DOI PubMed
47. Chen, T.; Guestrin, C. XGBoost: a scalable tree boosting system. *arXiv* **2016**, arXiv:1603.02754. Available online: https://doi.org/10.48550/arXiv.1603.02754. (accessed 2026-09-21).
48. Rodríguez-Pérez, R.; Bajorath, J. Evolution of support vector machine and regression modeling in chemoinformatics and drug discovery. *J. Comput. Aided. Mol. Des.* **2022**, *36*, 355-62. DOI PubMed PMC
49. Sakiyama, Y. The use of machine learning and nonlinear statistical tools for ADME prediction. *Expert. Opin. Drug. Metab. Toxicol.* **2009**, *5*, 149-69. DOI PubMed
50. Mansouri, K.; Grulke, C. M.; Judson, R. S.; Williams, A. J. OPERA models for predicting physicochemical properties and environmental fate endpoints. *J. Cheminform.* **2018**, *10*, 10. DOI PubMed PMC
51. Venkataraman, M.; Rao, G. C.; Madavareddi, J. K.; Maddi, S. R. Leveraging machine learning models in evaluating ADMET properties for drug discovery and development. *ADMET. DMPK.* **2025**, *13*, 2772. DOI PubMed PMC
52. Pantic, I.; Paunovic, J.; Cumic, J.; Valjarevic, S.; Petroianu, G. A.; Corridon, P. R. Artificial neural networks in contemporary toxicology research. *Chem. Biol. Interact.* **2023**, *369*, 110269. DOI PubMed
53. Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; Dahl, G. E. Neural message passing for quantum chemistry. *arXiv* **2017**, arXiv:1704.01212. Available online: https://doi.org/10.48550/arXiv.1704.01212. (accessed 2026-09-21).

54. Zhang, Z.; Tell, L. A.; Lin, Z. Development of machine learning and chemical language model-based QSAR models for predicting drug residue depletion half-lives in plasma and tissues of cattle across various administration routes. *J. Vet. Pharmacol. Ther.* **2026**, *49*, 150-71. DOI PubMed
55. Umer, M. S.; Nabeel, M.; Athar, U.; et al. Large language models meet molecules: a systematic review of advances and challenges in AI-driven cheminformatics. *Arch. Computat. Methods. Eng.* **2026**, *33*, 4867-908. DOI
56. Limbu, S.; Zakka, C.; Dakshanamurthy, S. Predicting dose-range chemical toxicity using novel hybrid deep machine-learning method. *Toxics* **2022**, *10*, 706. DOI PubMed PMC
57. Zhang, J.; Li, H.; Zhang, Y.; et al. Computational toxicology in drug discovery: applications of artificial intelligence in ADMET and toxicity prediction. *Brief. Bioinform.* **2025**, *26*, bbaf533. DOI PubMed PMC
58. Tropsha, A. Best practices for QSAR model development, validation, and exploitation. *Mol. Inform.* **2010**, *29*, 476-88. DOI PubMed
59. Wallach, I.; Heifets, A. Most ligand-based classification benchmarks reward memorization rather than generalization. *J. Chem. Inf. Model.* **2018**, *58*, 916-32. DOI PubMed
60. Sheridan, R. P. Time-split cross-validation as a method for estimating the goodness of prospective prediction. *J. Chem. Inf. Model.* **2013**, *53*, 783-90. DOI PubMed
61. OECD. Guidance document on the validation of (quantitative) structure-activity relationship [(Q)SAR] models. 2014. https://www.oecd.org/en/publications/guidance-document-on-the-validation-of-quantitative-structure-activity-relationship-q-sar-models_9789264085442-en.html. (accessed 2026-09-21).
62. Kar, S.; Roy, K.; Leszczynski, J. Applicability domain: a step toward confident predictions and decidability for QSAR modeling. *Methods. Mol. Biol.* **2018**, *1800*, 141-69. DOI PubMed
63. Alqahtani, S. Improving on in-silico prediction of oral drug bioavailability. *Expert. Opin. Drug. Metab. Toxicol.* **2023**, *19*, 665-70. DOI PubMed
64. Kumar, R.; Sharma, A.; Siddiqui, M. H.; Tiwari, R. K. Prediction of human intestinal absorption of compounds using artificial intelligence techniques. *Curr. Drug. Discov. Technol.* **2017**, *14*, 244-54. DOI PubMed
65. Kamiya, Y.; Omura, A.; Hayasaka, R.; et al. Prediction of permeability across intestinal cell monolayers for 219 disparate chemicals using in vitro experimental coefficients in a pH gradient system and in silico analyses by trivariate linear regressions and machine learning. *Biochem. Pharmacol.* **2021**, *192*, 114749. DOI PubMed
66. Wang, D.; Jin, J.; Shi, G.; et al. ADMET evaluation in drug discovery: 21. Application and industrial validation of machine learning algorithms for Caco-2 permeability prediction. *J. Cheminform.* **2025**, *17*, 3. DOI PubMed PMC
67. Rácz, A.; Vincze, A.; Volk, B.; Balogh, G. T. Extending the limitations in the prediction of PAMPA permeability with machine learning algorithms. *Eur. J. Pharm. Sci.* **2023**, *188*, 106514. DOI PubMed
68. Narita, I.; Todo, H.; Fujiwara, C.; et al. In silico model to predict dermal absorption of chemicals in finite dose conditions. *J. Toxicol. Sci.* **2025**, *50*, 171-86. DOI PubMed
69. Sarti, D.; Wagner, J.; Palma, F.; et al. Interpretable machine learning unveils key predictors and default values in an expanded database of human in vitro dermal absorption studies with pesticides. *Regul. Toxicol. Pharmacol.* **2025**, *159*, 105801. DOI PubMed
70. Chiu, Y. W.; Tung, C. W.; Wang, C. C. Multitask learning for predicting pulmonary absorption of chemicals. *Food. Chem. Toxicol.* **2024**, *185*, 114453. DOI PubMed
71. Simon, L. Advancing exposure science through artificial intelligence: neural ordinary differential equations for predicting blood concentrations of volatile organic compounds. *Ecotoxicol. Environ. Saf.* **2025**, *292*, 117928. DOI PubMed
72. Feschuk, A. M.; Law, R. M.; Maibach, H. I. Comparative efficacy of reactive skin decontamination lotion (RSDL): a systematic review. *Toxicol. Lett.* **2021**, *349*, 109-14. DOI PubMed
73. Dawson, D. E.; Ingle, B. L.; Phillips, K. A.; Nichols, J. W.; Wambaugh, J. F.; Tornero-Velez, R. Designing QSARs for parameters of high-throughput toxicokinetic models using open-source descriptors. *Environ. Sci. Technol.* **2021**, *55*, 6505-17. DOI PubMed PMC
74. Mi, K.; Chou, W. C.; Chen, Q.; et al. Predicting tissue distribution and tumor delivery of nanoparticles in mice using machine learning models. *J. Control. Release.* **2024**, *374*, 219-29. DOI PubMed PMC
75. Liu, L.; Zhang, L.; Feng, H.; et al. Prediction of the blood-brain barrier (BBB) permeability of chemicals based on machine-learning and ensemble methods. *Chem. Res. Toxicol.* **2021**, *34*, 1456-67. DOI PubMed
76. Nguyen, T.; Rana, M. M.; Mukta, F. T.; Zhan, C.; Nguyen, D. D. Geometric multi-color message passing graph neural networks for blood–brain barrier permeability prediction. *Mol. Syst. Des. Eng.* **2026**, *11*, 436-46. DOI
77. Huang, E. T. C.; Yang, J. S.; Liao, K. Y. K.; et al. Predicting blood-brain barrier permeability of molecules with a large language model and machine learning. *Sci. Rep.* **2024**, *14*, 15844. DOI PubMed PMC
78. Li, J.; Sun, X.; Xu, J.; Tan, H.; Zeng, E. Y.; Chen, D. Transplacental transfer of environmental chemicals: roles of molecular descriptors and placental transporters. *Environ. Sci. Technol.* **2021**, *55*, 519-28. DOI PubMed

79. Chen, X.; Yao, J.; Ma, Y.; et al. Rapid screening of chemicals with placental transfer risk using interpretable machine learning. *Environ. Sci. Technol. Lett.* **2024**, *11*, 798-804. DOI
80. Guan, R.; Cai, R.; Guo, B.; Wang, Y.; Zhao, C. A data-driven computational framework for assessing the risk of placental exposure to environmental chemicals. *Environ. Sci. Technol.* **2024**, *58*, 7770-81. DOI PubMed
81. Edbert Duru, C. Forever chemicals could expose the human fetus to xenobiotics by binding to placental enzymes: prescience from molecular docking, DFT, and machine learning. *Comput. Toxicol.* **2023**, *26*, 100274. DOI
82. Huang, X.; Chen, J.; Liu, P. Assessing chemical exposure risk in breastfeeding infants: an explainable machine learning model for human milk transfer prediction. *Ecotoxicol. Environ. Saf.* **2025**, *289*, 117707. DOI PubMed
83. Maeshima, T.; Yoshida, S.; Watanabe, M.; Itagaki, F. Prediction model for milk transfer of drugs by primarily evaluating the area under the curve using QSAR/QSPR. *Pharm. Res.* **2023**, *40*, 711-9. DOI PubMed PMC
84. Zhao, C.; Zhang, H.; Zhang, X.; et al. Prediction of milk/plasma drug concentration (M/P) ratio using support vector machine (SVM) method. *Pharm. Res.* **2006**, *23*, 41-8. DOI PubMed
85. Grant, N.; Machado Reyes, D.; Yang, Z.; Wan, L.; Wang, C.; Yan, P. Blood brain barrier permeability prediction with artificial intelligence and machine learning: a meta-review and future directions. *Discov. Artif. Intell.* **2025**, *5*, 494. DOI
86. Nabi, A. E.; Pouladvand, P.; Liu, L.; Hua, N.; Ayubcha, C. Machine learning in drug development for neurological diseases: a review of blood brain barrier permeability prediction models. *Mol. Inform.* **2025**, *44*, e202400325. DOI PubMed PMC
87. Agudelo-Pérez, S.; Botero-Rosas, D.; Rodríguez-Alvarado, L.; Espitia-Angel, J.; Raigoso-Díaz, L. Artificial intelligence applied to the study of human milk and breastfeeding: a scoping review. *Int. Breastfeed. J.* **2024**, *19*, 79. DOI PubMed PMC
88. Lai, Y.; Chu, X.; Di, L.; et al. Recent advances in the translation of drug metabolism and pharmacokinetics science for drug discovery and development. *Acta. Pharm. Sin. B.* **2022**, *12*, 2751-77. DOI PubMed PMC
89. Sun, H.; Veith, H.; Xia, M.; Austin, C. P.; Tice, R. R.; Huang, R. Prediction of cytochrome P450 profiles of environmental chemicals with QSAR models built from drug-like molecules. *Mol. Inform.* **2012**, *31*, 783-92. DOI PubMed PMC
90. Tyzack, J. D.; Mussa, H. Y.; Williamson, M. J.; Kirchmair, J.; Glen, R. C. Cytochrome P450 site of metabolism prediction from 2D topological fingerprints using GPU accelerated probabilistic classifiers. *J. Cheminform.* **2014**, *6*, 29. DOI PubMed PMC
91. Yang, H.; Liu, J.; Chen, K.; et al. D-CyPre: a machine learning-based tool for accurate prediction of human CYP450 enzyme metabolic sites. *PeerJ. Comput. Sci.* **2024**, *10*, e2040. DOI PubMed PMC
92. Ryu, J. Y.; Lee, J. H.; Lee, B. H.; Song, J. S.; Ahn, S.; Oh, K. S. PredMS: a random forest model for predicting metabolic stability of drug candidates in human liver microsomes. *Bioinformatics* **2022**, *38*, 364-8. DOI PubMed
93. Sasahara, K.; Shibata, M.; Sasabe, H.; et al. Predicting drug metabolism and pharmacokinetics features of in-house compounds by a hybrid machine-learning model. *Drug. Metab. Pharmacokinet.* **2021**, *39*, 100395. DOI PubMed
94. Kamiya, Y.; Handa, K.; Miura, T.; et al. Machine learning prediction of the three main input parameters of a simplified physiologically based pharmacokinetic model subsequently used to generate time-dependent plasma concentration data in humans after oral doses of 212 disparate chemicals. *Biol. Pharm. Bull.* **2022**, *45*, 124-8. DOI PubMed
95. Litsa, E. E.; Das, P.; Kavraki, L. E. Machine learning models in the prediction of drug metabolism: challenges and future perspectives. *Expert. Opin. Drug. Metab. Toxicol.* **2021**, *17*, 1245-7. DOI PubMed PMC
96. Tyzack, J. D.; Kirchmair, J. Computational methods and tools to predict cytochrome P450 metabolism for drug discovery. *Chem. Biol. Drug. Des.* **2019**, *93*, 377-86. DOI PubMed PMC
97. Wang, D.; Liu, W.; Shen, Z.; et al. Deep learning based drug metabolites prediction. *Front. Pharmacol.* **2019**, *10*, 1586. DOI PubMed PMC
98. Chao, P.; Uss, A. S.; Cheng, K. C. Use of intrinsic clearance for prediction of human hepatic clearance. *Expert. Opin. Drug. Metab. Toxicol.* **2010**, *6*, 189-98. DOI PubMed
99. Tran, T. T. V.; Tayara, H.; Chong, K. T. Artificial intelligence in drug metabolism and excretion prediction: recent advances, challenges, and future perspectives. *Pharmaceutics* **2023**, *15*, 1260. DOI PubMed PMC
100. Bois, F. Y.; Jamei, M.; Clewell, H. J. PBPK modelling of inter-individual variability in the pharmacokinetics of environmental chemicals. *Toxicology* **2010**, *278*, 256-67. DOI PubMed
101. Tucker, G. T. Measurement of the renal clearance of drugs. *Br. J. Clin. Pharmacol.* **1981**, *12*, 761-70. DOI PubMed PMC
102. Lee, W.; Kim, R. B. Transporters and renal drug elimination. *Annu. Rev. Pharmacol. Toxicol.* **2004**, *44*, 137-66. DOI PubMed
103. Wang, Z. J.; Yin, O. Q.; Tomlinson, B.; Chow, M. S. OCT2 polymorphisms and in-vivo renal functional consequence: studies with metformin and cimetidine. *Pharmacogenet. Genomics.* **2008**, *18*, 637-45. DOI PubMed
104. Ryu, S.; Yamaguchi, E.; Sadegh Modaresi, S. M.; et al. Evaluation of 14 PFAS for permeability and organic anion transporter interactions: implications for renal clearance in humans. *Chemosphere* **2024**, *361*, 142390. DOI PubMed PMC
105. Sharifi, M.; Ghafourian, T. Estimation of biliary excretion of foreign compounds using properties of molecular structure. *AAPS. J.* **2014**, *16*, 65-78. DOI PubMed PMC

106. Tátrai, P.; Erdő, F.; Krajcsi, P. Role of hepatocyte transporters in drug-induced liver injury (DILI)-in vitro testing. *Pharmaceutics* **2022**, *15*, 29. DOI PubMed PMC
107. Nakanishi, T.; Tamai, I. Interaction of drug or food with drug transporters in intestine and liver. *Curr. Drug. Metab.* **2015**, *16*, 753-64. DOI PubMed
108. Baker, M.; Parton, T. Kinetic determinants of hepatic clearance: plasma protein binding and hepatic uptake. *Xenobiotica* **2007**, *37*, 1110-34. DOI PubMed
109. Paine, S. W.; Barton, P.; Bird, J.; et al. A rapid computational filter for predicting the rate of human renal clearance. *J. Mol. Graph. Model.* **2010**, *29*, 529-37. DOI PubMed
110. Watanabe, R.; Ohashi, R.; Esaki, T.; et al. Development of an in silico prediction system of human renal excretion and clearance from chemical structure information incorporating fraction unbound in plasma as a descriptor. *Sci. Rep.* **2019**, *9*, 18782. DOI PubMed PMC
111. Hosey, C. M.; Broccatelli, F.; Benet, L. Z. Predicting when biliary excretion of parent drug is a major route of elimination in humans. *AAPS. J.* **2014**, *16*, 1085-96. DOI PubMed PMC
112. Chou, W. C.; Chen, Q.; Yuan, L.; et al. An artificial intelligence-assisted physiologically-based pharmacokinetic model to predict nanoparticle delivery to tumors in mice. *J. Control. Release.* **2023**, *361*, 53-63. DOI PubMed PMC
113. Obrezanova, O.; Martinsson, A.; Whitehead, T.; et al. Prediction of in vivo pharmacokinetic parameters and time-exposure curves in rats using machine learning from the chemical structure. *Mol. Pharm.* **2022**, *19*, 1488-504. DOI PubMed
114. Jia, X.; Wang, T.; Zhu, H. Advancing computational toxicology by interpretable machine learning. *Environ. Sci. Technol.* **2023**, *57*, 17690-706. DOI PubMed PMC
115. Tjoa, E.; Guan, C. A survey on explainable artificial intelligence (XAI): toward medical XAI. *IEEE. Trans. Neural. Netw. Learn. Syst.* **2021**, *32*, 4793-813. DOI PubMed
116. Hassija, V.; Chamola, V.; Mahapatra, A.; et al. Interpreting black-box models: a review on explainable artificial intelligence. *Cogn. Comput.* **2024**, *16*, 45-74. DOI
117. Wu, K.; Li, X.; Zhou, Z.; et al. Predicting pharmacodynamic effects through early drug discovery with artificial intelligence-physiologically based pharmacokinetic (AI-PBPK) modelling. *Front. Pharmacol.* **2024**, *15*, 1330855. DOI PubMed PMC
118. Wang, W.; Wang, N.; Wu, Y.; et al. An integrated AI-PBPK platform for predicting drug in vivo fate and tissue distribution in human and inter-species extrapolation. *Clin. Pharmacol. Ther.* **2025**, *118*, 865-75. DOI PubMed
119. Kamiya, Y.; Otsuka, S.; Miura, T.; et al. Plasma and hepatic concentrations of chemicals after virtual oral administrations extrapolated using rat plasma data and simple physiologically based pharmacokinetic models. *Chem. Res. Toxicol.* **2019**, *32*, 211-8. DOI PubMed