

Review

Artificial Intelligence-Driven Drug Toxicity Prediction: Advances, Challenges, and Future Directions

Ruiqiu Zhang ^{1,2}, Hairuo Wen ^{2,3}, Zhi Lin ^{2,3}, Bo Li ^{1,2} and Xiaobing Zhou ^{2,3,*}¹ National Institutes for Food and Drug Control, Chinese Academy of Medical-Sciences and Peking Union Medical College, Beijing 100730, China² National Center for Safety Evaluation of Drugs, National Institutes for Food and Drug Control, Beijing 100076, China³ State Key Laboratory of Drug Regulatory Science, Beijing 102629, China

* Correspondence: zhxb@nifdc.org.cn

Abstract

Drug toxicity prediction plays a crucial role in the drug research and development process, ensuring clinical drug safety. However, traditional methods are hampered by high cost, low throughput, and uncertainty of cross-species extrapolation, which has become a key bottleneck restricting the efficiency of new drug research and development. The breakthrough development of Artificial Intelligence (AI) technology, especially the application of deep learning and multimodal data fusion strategy, is reshaping the scientific paradigm of drug toxicology assessment. In this review, we focus on the application of AI in the field of drug toxicity prediction and systematically summarize the relevant literature and development status globally in the past years. The application of various toxicity databases in the prediction was elaborated in detail, and the research results and methods for the prediction of different toxicity endpoints were analyzed in depth, including acute toxicity, carcinogenicity, organ-specific toxicity, etc. Furthermore, this paper discusses the application progress of AI technologies (e.g., machine learning and deep learning model) in drug toxicity prediction, analyzes their advantages and challenges, and outlines the future development direction. It aims to provide a comprehensive and in-depth theoretical framework and actionable technical strategies for toxicity prediction in drug development.

Keywords: Artificial Intelligence; machine learning; deep learning; toxicity endpoints; transfer learning



Academic Editor: Matthew J. Winter

Received: 19 May 2025

Revised: 16 June 2025

Accepted: 19 June 2025

Published: 23 June 2025

Citation: Zhang, R.; Wen, H.; Lin, Z.; Li, B.; Zhou, X. Artificial Intelligence-Driven Drug Toxicity Prediction: Advances, Challenges, and Future Directions. *Toxics* **2025**, *13*, 525. <https://doi.org/10.3390/toxics13070525>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Drugs, as chemical substances that can change the physiological or psychological state of organisms, can not only be used for disease prevention, diagnosis, and treatment, improve the quality of life of patients, and even extend the length of human life but also may have different degrees of toxic effects on the body, cause some adverse reactions, cause permanent damage to specific organs, and even cause disability or death of patients. Therefore, determining how to avoid the safety risks of drugs is one of the main challenges faced by drug development, and it is also a crucial and urgent public health problem worldwide [1,2]. In the process of drug research and development, drug toxicity prediction is the key link to ensure drug safety. The toxicity of drugs is not only related to the life and health of patients but also has a decisive impact on the cost, cycle, and marketing approval of drug research and development. Statistical analysis found that among the reasons for the failure of drug research and development, safety factors such as toxicity and side effects

accounted for 30% [3,4], making them the main reason for the failure of drug research and development beyond pharmacodynamic factors. Some of the drugs already on the market have also been withdrawn due to major safety problems. The failure of listed drugs will not only lead to huge economic losses but also endanger human health. Therefore, in the early stages of drug research and development, the success rate of innovative drug research and development can be improved by accurately evaluating and predicting the safety of candidate compounds, removing toxic compounds from the candidate compound library, or carrying out targeted structural transformation of toxic compounds.

Traditional methods of drug toxicity prediction are mainly based on in vitro experiments and animal experiments [2,3]. These methods have many limitations, such as long experimental cycles, high costs, and limited prediction accuracy [3]. Although animal experiments can simulate human physiological responses to a certain extent, there are species differences, and the accuracy of extrapolation of experimental results to the human body is limited. Traditional drug toxicity prediction methods can struggle to meet the needs of modern drug research and development. Therefore, AI-enabled drug toxicity prediction technology came into being [5,6], which has brought new opportunities to solve this problem. AI technology, especially machine learning and deep learning algorithms [7], has powerful data processing and pattern recognition capabilities. In drug toxicity prediction, AI can quickly analyze massive data of drug structure, activity, toxicity, etc., and mine the hidden rules and associations so as to establish a high-precision prediction model. Compared with traditional methods, AI technology has higher efficiency and accuracy and can quickly screen potentially toxic drugs in the early stage of drug development, reducing the cost and risk of research and development. At the same time, AI models can also be optimized and updated by constantly learning new data, improving the prediction performance, and providing more reliable support for drug toxicity prediction [1,8].

This review focuses on the application of AI in the field of drug toxicity prediction and systematically combs the relevant literature and development at home and abroad in recent years. The application of various toxicity databases in the prediction was elaborated in detail, and the research results and methods for the prediction of different toxicity endpoints were analyzed in depth, including acute toxicity, carcinogenicity, organ-specific toxicity, etc. At the same time, this paper discusses the application progress of AI technology, such as machine learning and deep learning models, in drug toxicity prediction, analyzes its advantages and challenges, and looks forward to future development directions, aiming to provide a comprehensive and in-depth theoretical reference and technical guidance for toxicity prediction in the process of drug research and development.

2. Drug Toxicity Prediction: Databases and Tools

2.1. Toxicity Database

2.1.1. Toxicology Resources for Intelligent Computation

The Toxicology Resources for Intelligent Computation (TOXRIC) database is a comprehensive toxicity database, containing a large number of toxicity data of compounds, which are derived from a variety of experiments and literature. Its data cover acute toxicity, chronic toxicity, carcinogenicity, and other types of toxicity, and the species involved include humans, animals, and aquatic organisms [9]. In the prediction of drug toxicity, the TOXRIC database provides rich training data from which researchers can extract the structure of drug molecules and the corresponding toxicity information for the construction and training of machine learning models.

2.1.2. Integrated Chemical Environment (ICE)

The ICE database integrates chemical substance information and toxicity data from multiple data sources [10], including the basic properties of chemical substances (such as molecular formula, structural formula, etc.), toxicological data (such as median lethal dose (LD50), half maximal inhibitory concentration (IC50), etc.) [11], and environmental fate information. The data in this database have high quality and reliability and provide researchers with comprehensive chemical information and toxicity references in drug toxicity prediction research.

2.1.3. Distributed Structure-Searchable Toxicity (DSSTox) Database

The DSSTox database is a large searchable toxicity database that contains a large number of structure, toxicity, and related experimental data on chemical substances [12]. Toxval is part of DSSTox and provides standardized data on toxicity values. These data are widely used in drug toxicity prediction, environmental risk assessment, and other fields. In drug research and development, researchers can use the data in DSSTox and Toxval databases to preliminarily evaluate and screen the toxicity of drug molecules.

2.1.4. DrugBank

The DrugBank database is a comprehensive and freely accessible online database, which provides detailed information on drugs and their targets, including basic data, chemical structure, pharmacological data on drugs, and the sequence, structure, and pathway information of targets [13]. In addition, it also contains clinical information such as clinical trials, adverse reactions, and drug interactions, and provides functions such as external links, data downloads, and Application Programming Interface (API) access.

2.1.5. ChEMBL

The ChEMBL database is a manually curated database of bioactive molecules with drug-like properties. It brings together chemical, bioactive, and genomic data [14]. It aims to translate genomic information into effective new drugs. The database integrates a large amount of data from authoritative journals, patents, and laboratory research reports, providing a wealth of compound structure information, bioactivity data, drug target information, and absorption, distribution, metabolism, excretion, and toxicity (ADMET) data. Researchers can access and query data in many ways, including chemical structure searches, drug name searches, target searches, etc., and use the analysis tools provided by the database for activity clustering, structural similarity searches, and other analyses.

2.1.6. Online Chemical Modeling Environment (OCHEM)

The OCHEM database currently contains 4,064,457 records with 695 attributes, sourced from 20,932 distinct references [15]. Through OCHEM, users can build Quantitative Structure–Activity Relationship (QSAR) models to predict chemical properties or screen chemical libraries according to structural alarms of mutagenicity, skin sensitization, aquatic toxicity, and other endpoints.

2.1.7. PubChem

PubChem is a world-renowned database of chemical substances, which contains massive data on the structure, activity, and toxicity of chemical substances [16]. It integrates information from scientific research literature, experimental reports, and other databases, and has the characteristics of a large amount of data and timely updates. In drug toxicity prediction, the PubChem database is one of the important data sources, from which researchers can obtain a large number of drug molecular data and corresponding toxicity information for model training and validation.

In addition to the seven toxicity databases mentioned above, Table 1 also lists a variety of existing databases related to drug toxicology resources, showing the basic information, data scale, and characteristics of the database. These databases provide all kinds of information about the toxicity of drugs or exogenous substances to cells, organs, or the whole body. Toxicity endpoints include mutagenicity, carcinogenicity, skin sensitization, teratogenicity, endocrine interference, organ-specific toxicity, and other toxicity endpoints. Data types are not only limited to in vivo and in vitro test data but also include predictive data QSAR models.

2.2. Biological Experimental Data

2.2.1. In Vitro Cytotoxicity Test Data

The in vitro cytotoxicity test is a common method to evaluate the toxicity of drugs. The toxicity information of drugs at the cellular level can be obtained by detecting the growth inhibition, apoptosis, and other indicators of drugs on cells. These experimental data can be used as the training data of the AI model to construct the cytotoxicity prediction model. Common cytotoxic tests in vitro include the MTT test [17], the CCK-8 test [18], etc.

2.2.2. Animal Experiment Data

Although animal experiments have some limitations in the prediction of drug toxicity, it is still one of the important means to evaluate drug toxicity. Animal experimental data include pharmacokinetic parameters, tissue and organ toxicity, reproductive toxicity, and other information on drugs in animals. These data can provide more comprehensive toxicity information for AI models and help to improve the prediction accuracy of the model. However, when using animal experimental data, attention should be paid to the transformation and application of data to reduce the impact of species differences on the prediction results.

2.3. Clinical Data

2.3.1. Food and Drug Administration (FDA) Adverse Event Reporting System

The FDA Adverse Event Reporting System (FAERS) is an adverse drug reaction reporting system established by the FDA [19]. It has collected a large number of adverse drug reaction reports from the drug market. These reports contain information on various toxic reactions of drugs during clinical use, such as drug-induced liver injury, kidney injury, cardiovascular toxicity, etc. By mining FAERS data, we can build a drug toxicity prediction model based on clinical data and provide an important reference for drug safety evaluation.

2.3.2. Electronic Medical Record System

The electronic medical record system (EMR) records the detailed medical information of patients, including drug use records, disease diagnosis, laboratory test results, etc. These data can be used to analyze the toxicity of drugs in practical clinical applications and provide more authentic and comprehensive data support for drug toxicity prediction. However, there are some problems with the quality and integrity of EMR data. It mainly includes the following aspects: (1) inconsistent formats (for example, mixed date/number representations leading to parsing errors), (2) missing key fields (approximately 30% dose omissions, 41% unrecorded concomitant drugs, and 25–60% missing organ toxicity biomarkers), and (3) non-standardized terms. To solve these problems, an integrated preprocessing pipeline is required—standardizing the format through regular expressions, inputting missing values using the k-nearest neighbor algorithm, mapping free text to the standardized ontology through natural language processing (NLP) techniques [20], and filtering out biologically unreliable values—in order to improve the reliability of computational toxicology applications.

Table 1. Databases and tools used for chemical toxicity prediction.

Number	Database Name	Database Content	Characteristic	Developer	Website
1	Tox21 [21]	It contains data on more than 10,000 chemicals, as well as nuclear receptors and stress response pathways, generating more than 150 million data points.	High throughput screening data; Multiple toxicological toxicity endpoints; Data visualization and open access	FDA	https://tripod.nih.gov/tox21/challenge/
2	Comp Tox Dashboard [22]	It contains data on more than 1.2 million chemicals and provides information about chemical structure, environmental behavior, and biological activity.	Real-time prediction; Batch Search; Advanced search; Data integration	Environmental Protection Agency (EPA)	https://comptox.epa.gov/dashboard
3	Toxicity Reference Database (ToxRefDB) [23]	It contains in vivo research data of more than 5900 guideline or class guideline studies from more than 1100 chemicals and provides quantitative dose-response data of each dose treatment group, including data of the control group, as well as dose, effect value, and variance information.	High-quality data; All-purpose; Data integration; Distinguishing between missing and negative endpoints	EPA	https://www.epa.gov/chemical-research/downloadable-computational-toxicology-data
4	Toxin and Toxin Target Database (T3DB) [24]	It contains 3678 toxins, described by 41,602 synonyms, including pollutants, pesticides, drugs, and food toxins. These toxins were associated with 2073 corresponding toxin target records, with a total of 42,374 toxin target associations.	Data diversity and accessibility; Potential applications; Data integration	The Chinese University of Hong Kong	http://www.t3db.ca/
5	Therapeutic Target Database (TTD) [25]	TTD contains more than 3500 drug targets and nearly 40,000 drug molecules. The database provides information about target-related diseases and helps researchers understand the potential role of targets in disease treatment.	Reliable data source; Clear target classification; Powerful retrieval function; Rich auxiliary functions	National University of Singapore	http://db.idrblab.net/ttd/
6	Side Effect Resource (SIDER) [26]	It contains information on 1430 drugs and 5868 side effects, covering a variety of treatment areas. The database provides frequency data of each side effect, allowing users to assess the possibility of specific adverse reactions.	Comprehensiveness; Reliable data source; Multilingual support; Structured data	Technical University of Berlin	http://sideeffects.embl.de/
7	The Marker [27]	It contains 218 efficacy biomarkers, 624 safety biomarkers, 104 monitoring biomarkers, 15,893 predictive biomarkers, and 103 alternative endpoints. These data cover a large number of drugs and a wide range of disease categories, not limited to anti-cancer therapy.	Systematic organization; Rich data; User-friendliness; Free access	IDRBLAB	https://idrblab.org/themarker

Table 1. Cont.

Number	Database Name	Database Content	Characteristic	Developer	Website
8	Comparative Toxicogenomics Database (CTD) [28]	CTD contains the interaction information between chemical substances and genes/proteins, as well as the interaction information between chemical substances and phenotypes.	User-friendly interface; The database is searchable, accessible, interoperable, and reusable; Non-redundant data; Comparative analysis is available	North Carolina State University	https://ctdbase.org
9	Gene Expression Omnibus (GEO) [29]	GEO's significance for ML models stems from its vast repository of gene expression data, encompassing both microarray and RNA-seq datasets. These datasets enable the training of models to predict how chemicals or drugs alter gene expression, thereby providing insights into their potential molecular-level toxic effects.	High data quality; Various forms of data storage; Data visualization; Easy data retrieval; Timely data update	National Center for Biotechnology Information (NCBI)	https://www.ncbi.nlm.nih.gov/geo/
10	Drug Matrix [30]	It contains the comprehensive results of thousands of highly controlled and standardized toxicological experiments involving rat or primary rat hepatocytes, which are systematically treated with therapeutic, industrial, and environmental chemicals, including non-toxic doses and toxic doses	High-quality data; Multiple data types; User-friendliness; Data analysis tools	National Toxicology Program (NTP)	https://norecopa.no/3r-guide/drugmatrix https://ntp.niehs.nih.gov/drugmatrix/index.html
11	Kyoto Encyclopedia of Genes and Genomes (KEGG) [31]	KEGG is an encyclopedia of genes and genomes, which offers structured information on metabolic pathways, molecular interactions, and gene functions.	Highly organized data structure; Multiple data types; Cross-species comparison; Multiple applications	Kanehisa Laboratory	https://www.genome.jp/kegg
12	Universal Protein Database (UniProt) [32]	It is a comprehensive database of protein sequence and function information. It provides detailed annotation of protein function, including protein function, subcellular localization, post-translational modification, protein-protein interaction, and pathway information.	Functional notes; Data consolidation; Sequence similarity search; Proteomic analysis	European Bioinformatics Institute (EBI)	https://www.uniprot.org/
13	Therapeutics Data Commons (TDC) [33]	TDC covers a wide range of learning tasks, including target discovery, activity screening, efficacy, safety, and manufacturing, involving small molecules, antibodies, vaccines, and other biomedical products.	Three-tier structure; Rich datasets; Machine learning tasks; Data processing and evaluation	Harvard University	https://tdcommons.ai/
14	Toxbank [34]	Dedicated database for toxicity data management and modeling repository of “gold” compounds and selected test compounds, as well as reference resources of cells, cell lines, and tissues related to in vitro toxicity studies.	Data management; Interdisciplinary cooperation; Data sharing	EU FP7 project	https://toxbank.net/

Table 1. Cont.

Number	Database Name	Database Content	Characteristic	Developer	Website
15	LiverTox: Clinical and Research Information on Drug-Induced Liver Injury (LIVERTOx) [35]	LiverTox provides the clinical characteristics of drug-induced liver injury, disease severity classification, and causality evaluation scale. The database contains detailed records of more than 1400 drugs, herbs, and dietary supplements. The records of each drug include background information, hepatotoxicity description, case reports, references, etc.	Free use; Timely update; Comprehensiveness; Clinical and research support	National Library of Medicine (NIH)	https://livertox.nih.gov/
16	Gene Expression Nebulas (GEN) [36]	Gen database integrates 323 high-quality transcriptome datasets, covering 50,500 samples of 30 species and 15,540,169 cells, and provides transcriptional maps under six biological scenarios: baseline reference, genetics, phenotype, environment, time, and space.	Data quality control and standardization; Analysis and visualization tools; User-friendliness	National Biological Information Center	https://ngdc.cncb.ac.cn/gen
17	The Cancer Genome Atlas (TCGA) [37]	The TCGA project conducted molecular characterization analysis on more than 20,000 primary cancers and matched normal samples, covering 33 cancer categories. It also contains detailed clinical data, such as patient survival time, treatment response, etc. These data are of great value for the study of clinical characteristics and treatment effects of cancer	Openness and sharing of data; Diversity and comprehensiveness of data; High quality and standardization of data	National Cancer Institute (NCI)	https://www.cancer.gov/ccg/research/genome-sequencing/tcga
18	GeneCards [38]	Genecards provide detailed information about all known and predicted human genes, including genome location, function, expression pattern, genetic variation, clinical relevance, and functional annotation of genes. It also integrates a variety of biological pathway information, providing the role and relationship of genes in different biological pathways	Data consolidation; User-friendliness; Analytical tools; Data access and mining	Weizmann Institute of Science	https://www.genecards.org/

3. Application of Artificial Intelligence and Core Algorithm Technology

3.1. Machine Learning Algorithm

3.1.1. Traditional Machine Learning Algorithm

The traditional machine learning (ML) algorithm is a kind of machine learning method used before the rise of deep learning (neural networks). These algorithms usually rely on feature engineering and statistical techniques rather than end-to-end learning methods of deep neural networks. We list three common traditional machine learning algorithms and show their schematic diagrams, as shown in Figure 1.

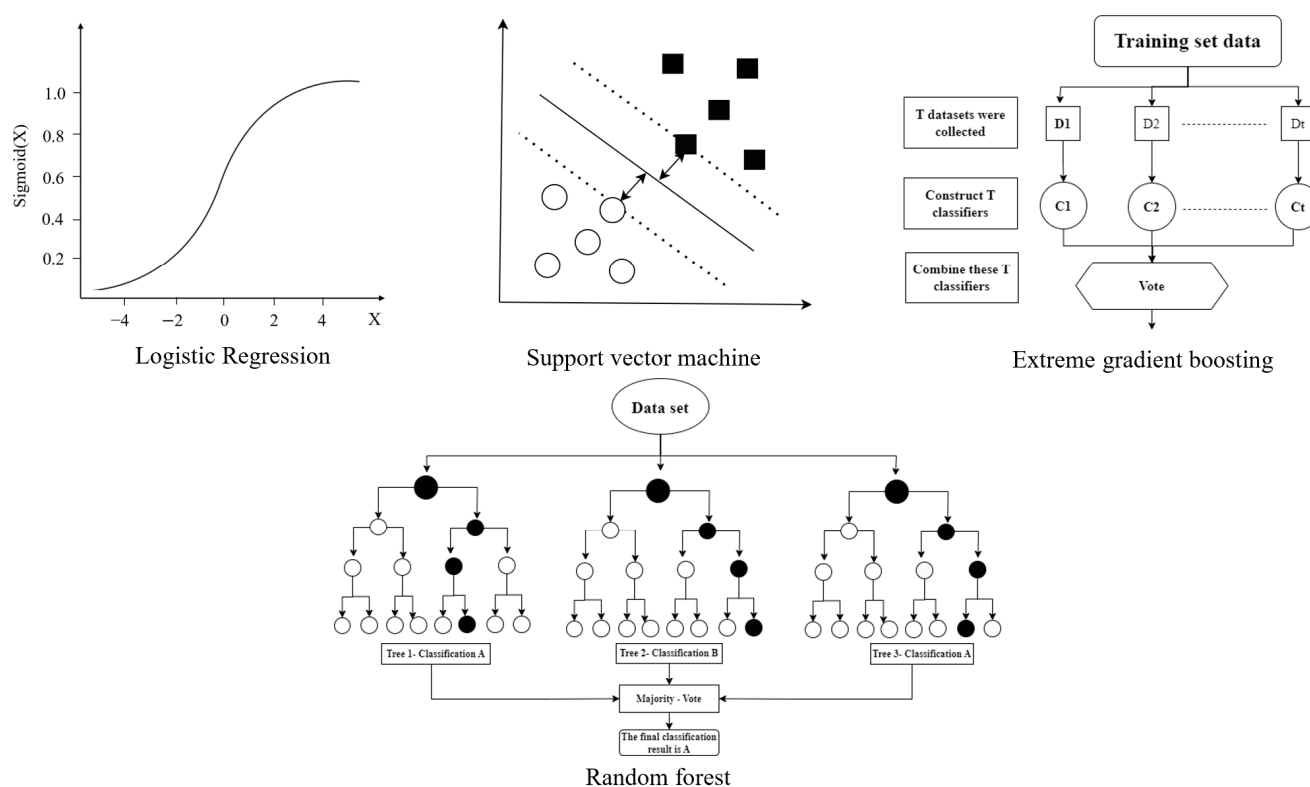


Figure 1. ML algorithm workflow for drug toxicity prediction.

Logistic regression (LR) is one of the most basic methods in the field of machine learning, and it is also a traditional model often used to interpret clinical trial data in biomedical research [39]. In the prediction of drug toxicity by logistic regression, the characteristics of drugs (molecular descriptors, physical and chemical properties, etc.) are associated with toxicity labels by constructing a linear model to calculate the probability of drug toxicity, which directly shows the weights of chemical descriptors in toxicity predictions [40]. Its advantage is that the model is simple, easy to understand and explain, and can directly display the weight of each feature in toxicity prediction, which has strong interpretability [41]. When predicting the acute toxicity of drugs, the linear relationship between some molecular structure characteristics and acute toxicity can be determined by the LR model. However, LR requires high linear separability of data and has limited ability to model complex nonlinear relationships [42].

Random forest (RF) is an integrated learning method [43]. It is based on multiple independent decision trees and obtains the final output by combining the prediction results of each tree. The algorithm can be used in feature selection, classification tasks, regression analysis, and other fields. At present, the random forest algorithm has been applied to many related works on toxicity prediction, such as the prediction of common side effects of

chemotherapy drugs [44], the probability of liver toxicity induced by inhibitors [45], and the risk assessment of drug toxicity [46].

The support vector machine (SVM) is a supervised learning algorithm based on the application of statistical theory [47], which is widely used in classification and regression problems. SVM separates different categories of data by finding an optimal classification hyperplane. The SVM algorithm is used to predict liver injury [48], kidney injury [49], cardiotoxicity [50], immunotoxicity [51], and bone marrow toxicity [52] caused by drugs and their metabolites. The experimental results show that SVM has excellent classification ability in processing small sample toxicity data and has good generalization performance.

Extreme gradient boosting (XGB) is an efficient gradient-lifting decision tree algorithm [53]. In the prediction of drug toxicity, it trains multiple decision trees iteratively to continuously fit the residuals and improve the prediction accuracy of the model. XGB has the advantages of fast computing speed and strong scalability and can handle large-scale datasets. However, the XGB model is relatively complex and struggles to adjust parameters, which requires some experience and skills.

3.1.2. Deep Learning Algorithm

Deep learning (DL) is a special subset of neural networks [54]. It forms a large-scale neural network with multiple hidden layers by deepening the depth of the network and uses the deep neural network (DNN) model to learn. It often supports the processing of thousands of objects and can be competent for more complex work. Raies et al. [55] used three multi-label classification methods (binary association, classifier chain, and unique multi-class) to predict multiple toxicity endpoints of the same compound. The comparative calculation results showed that the classifier chain algorithm had better performance in predicting multiple toxicity effects of drugs. Zhang et al. [56] established a multi-label k-nearest neighbor prediction method based on feature selection combined with multiple algorithms on a dataset containing 1080 drugs and 2260 side effects and integrated the basic classifiers according to the weighted strategy to obtain the final high-precision integration model. Kim et al. [57] analyzed the diversity of pathological findings, collected 21 pathological changes related to 41 drugs, and constructed an integrated model for multi-organ toxicity prediction based on gene expression data. The integrated model is composed of two parts: a k-nearest neighbor classification (KNN) classifier established separately for each pathology and the integrated model of liver or kidney formed by weighted calculation according to the correlation degree between pathology. On the dataset of liver and kidney, the cross-validation of the performance of the integrated model in predicting pathological results showed that the average area under the curve of all single pathological prediction models was 0.68, while the corresponding average of the integrated model was 0.88. In general, the effect of the integrated toxicity prediction model is better than that of the single pathological prediction model.

From this point of view, the development of deep learning algorithms and various interpretable framework algorithms are widely used in drug toxicity and prediction. Next, we will summarize the prediction of specific toxicity endpoints.

3.2. Molecular Representation

3.2.1. Molecular Fingerprint

The Morgan fingerprint is a kind of molecular fingerprint based on a circular substructure [58]. It generates substructures by gradually expanding the radius with each atom as the center on the molecular graph and encodes these substructures. Morgan fingerprint can effectively describe the local chemical structure characteristics of molecules and is widely used in similarity searches and model construction in drug toxicity prediction. For example,

the potential toxicity of drugs can be preliminarily determined by calculating the Morgan fingerprint similarity between drug molecules and known toxic molecules.

The Molecular Access System (MACCS) fingerprint is a molecular fingerprint based on specific atomic types and substructures. It contains 166 fixed substructure features and has certain chemical significance. In the prediction of drug toxicity, the MACCS fingerprint can be used as a molecular feature representation for model training and prediction. Due to its fixed characteristics, the computational speed is fast, but it may not be able to fully describe the complex structure information of molecules.

The Rdkit fingerprint is a molecular fingerprint implemented in the Rdkit chemical informatics toolkit. It combines a multi-seed structure and topological features and has good versatility and performance. Rdkit fingerprints are widely used in drug toxicity prediction, which can provide rich molecular structure information for the model and help to improve prediction accuracy.

3.2.2. Molecular Descriptor

Molecular descriptors are parameters that quantitatively describe various properties of molecules through mathematical methods, including physical and chemical properties (such as the molecular weight, number of hydrogen bond donors and receptors, lipid water partition coefficient, etc.), topological properties (such as the molecular connectivity index, path length, etc.), and electronic properties (such as charge distribution, frontier orbital energy, etc.). In the prediction of drug toxicity, molecular descriptors can reflect the characteristics of drug molecules from multiple perspectives. As the input of the machine learning model, molecular descriptors can help the model learn the relationship between molecular structure and toxicity.

3.2.3. Molecular Diagram

Molecular graphs represent drug molecules in a graph structure, in which nodes represent atoms and edges represent chemical bonds. Each node and edge can carry attribute information such as atomic type and bond type. Molecular graphs can intuitively display the structure and topology information of drug molecules. In deep learning models, such as the graph neural network (GNN) and gat, molecular graphs can be directly processed and learned by the model as input data. Through the analysis of molecular graphs, the model can capture the interaction and spatial relationship between atoms in molecules so as to predict drug toxicity more accurately. For example, in the GNN-based drug toxicity prediction model, the node features and edge features of the molecular graph are used as the input of the model. After the convolution operation of the multilayer graph, the potential representation of molecules is learned for toxicity prediction. We compared the performance of various molecular representation methods, as shown in Table 2.

Table 2. Molecular representation methods for drug toxicity prediction.

Category	Method	Core Principle	Key Features
Molecular Fingerprint	Morgan Fingerprint	Circular substructure expansion around each atom	- Radius-based atom environments - Encodes local chemical patterns
	MACCS Fingerprint	166 predefined substructure keys	- Fixed binary features - Chemically meaningful fragments (e.g., aromatic rings)
	RDKit Fingerprint	Hybrid of path-based and topological patterns	- Customizable bit length - Combines structural/topological diversity

Table 2. Cont.

Category	Method	Core Principle	Key Features
Molecular Descriptor	Physicochemical	Mathematical quantification of properties	- MW, LogP, H-bond donors/acceptors - Polar surface area
	Topological	Graph-theory indices	- Connectivity indices - Path lengths, branching degrees
	Electronic	Quantum chemical properties	- Partial charges - Frontier orbital energies
Molecular Graph	GNN	Atoms = Nodes, Bonds = Edges	- Node attributes: Atom type, charge - Edge attributes: Bond order, stereo

4. Research Progress of Different Toxicity Prediction

4.1. Acute Toxicity Prediction

4.1.1. Definition and Evaluation Index of Acute Toxicity

Acute toxicity describes the harmful effects of a substance that occur shortly after exposure, typically within 24 h [59]. It is commonly evaluated using the LD50, a standard toxicological measure. The LD50 represents the dose of a substance required to kill 50% of a test animal population during a specified observation period. It is a basic measure of the acute toxicity of drugs and plays an important role in drug classification, label management, and preliminary toxicity assessment. According to the regulations of the World Health Organization (WHO), drugs with an LD50 value of less than 5mg/kg are classified as highly toxic drugs, and their safety needs special attention. Common acute toxicity datasets comprise TOXRIC, ICE, EPA's DSSTox and ToxVal, ChemIDplus, eChemPortal, Japan's NITE (Chemical Risk Information Platform (CHRIP)), PubChem, and researcher-compiled custom datasets.

4.1.2. Relevant Research Results and Methods

The quality and size of acute toxicity datasets are often different, which will affect the robustness of prediction models. Lou et al. integrated the experimental data of 1735 cases of acute toxicity based on rabbits and 1679 cases of acute toxicity based on rats [60] and used machine learning to build a cross-species toxicity prediction model. This study found that in the model based on rat data training, the RF algorithm performed best, and its prediction accuracy was significantly better than that of SVM, XGB, and other control models. In view of the quality differences in the dataset, the research team proposed a two-stage optimization scheme: (1) data screening: eliminating abnormal samples through outlier detection (such as isolation forest) and consistency tests; (2) data enhancement: random transformation of the molecular structure (such as smile enumeration) and feature space expansion (such as molecular descriptor interpolation) technology are adopted to improve data diversity and model generalization ability. This study solves the problem of data heterogeneity in acute toxicity prediction through the strategy of combining data-driven and mechanism interpretation and provides a high-precision and interpretable calculation tool for drug safety assessment. However, due to the need to train multiple basic models, the computational complexity of this model is high, especially when dealing with large-scale datasets. The algorithm requires high data quality and needs data preprocessing and filtering to avoid the influence of noise and outliers. Li et al. developed an integrated learning model [61] based on a super learner to predict the acute toxicity of chemicals to rats. This study integrated the LD50 data of 9843 compounds and constructed 16 meta models using 4 molecular descriptors and a machine learning algorithm. The consistency model achieved R^2 values of 0.61 (five-fold cross-validation) and 0.64 (test set), with corresponding root mean square errors (RMSEs) of 0.55 and 0.64, respectively. RMSE is a standard metric

used to quantify the accuracy of predictive models. It measures the average magnitude of prediction errors between observed values and model-predicted values [62]. These results significantly outperform those of any single benchmark model. By integrating multiple base models, the consistency model enhances both the generalization capability and prediction accuracy. At the same time, it can flexibly select and adjust the basic model according to different datasets and problems. Feitosa et al. developed a machine learning tool called cyto-safe for the early identification of cytotoxic compounds [63]. In this study, the near-miss V3 undersampling method was used to balance the number of toxic and non-toxic samples, the ecfp4 molecular fingerprint was used to represent the structure of the compound, and the light Gradient-Boosting Machine (GBM) algorithm was used to build the model. Under a 1:5 undersampling ratio, the model performed best, with a sensitivity of 83%, and achieved high accuracy in the identification of cytotoxic compounds. It also constructs an interpretable prediction framework, which visually shows the relationship between molecular structure and toxicity prediction through a thermodynamic diagram. However, the dataset was limited to 3T3 and HEK 293 cell lines and did not cover other types of cells.

ML models can extract features from a variety of data sources and transform them into numerical information that can be used for toxicity prediction. Research shows that the RF model performs well in acute toxicity prediction and can effectively deal with complex toxicity data. DL, as a sub-field of machine learning, has surpassed traditional machine learning methods in the field of toxicity prediction because it can automatically construct complex features. For example, graph convolutional networks (GCNs) show stronger applicability in processing complex protein interaction network data.

4.2. Organ-Specific Toxicity Prediction

Organ-specific toxicity is critical, as drugs often target particular tissues or organs, causing adverse reactions due to local accumulation or organ-specific sensitivity. Drug-induced organ toxicity encompasses hepatotoxicity, cardiotoxicity, nephrotoxicity, respiratory toxicity, and neurotoxicity, as shown in Figure 2.

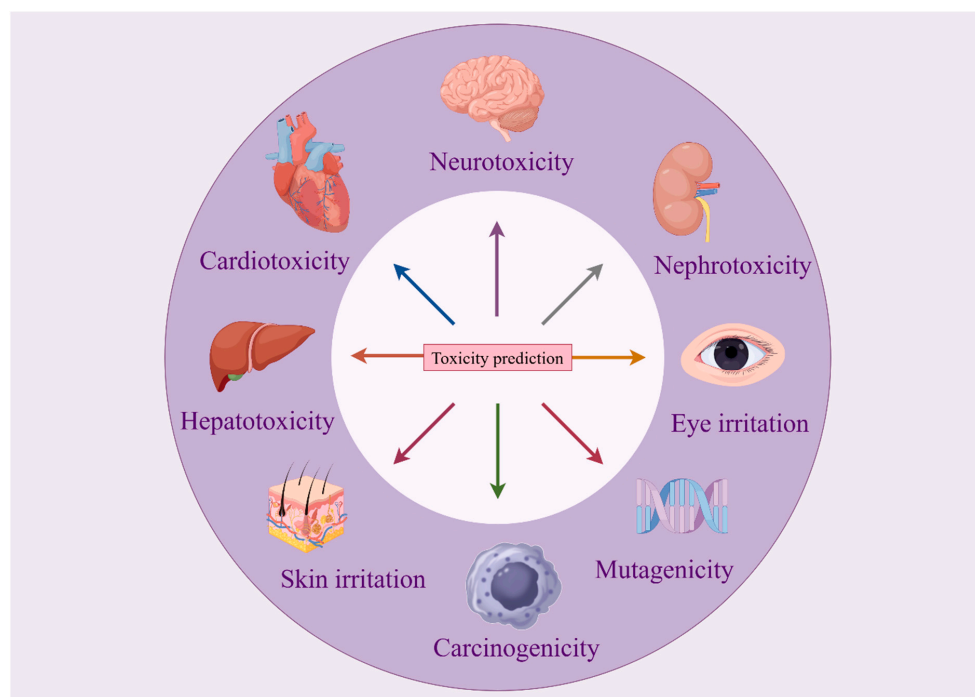


Figure 2. Schematic diagram of drug organ toxicity.

4.2.1. Hepatotoxicity

The liver is the main organ of drug metabolism, and hepatotoxicity is one of the common adverse reactions of drugs. The mechanisms of hepatotoxicity include drug-induced liver injury (DILI), interference with liver metabolic function, and immune-mediated liver injury. Hepatotoxicity is one of the most common and serious drug toxicity. The accurate prediction of drug hepatotoxicity is very important to ensure the safety of patients. AI technology has been widely used in the prediction of liver toxicity. By analyzing the structural characteristics of drug molecules, metabolic pathways, and the interaction with liver-related targets, the prediction model is constructed. Chen et al. developed ResNet18DNN, a deep neural network utilizing composite structural images to enhance hepatotoxicity prediction performance [64]. Separately, Füzi et al. proposed a systems biology approach integrating biological pathway-related descriptors, demonstrating superior predictive capability compared to traditional machine learning models such as Gradient-Boosting Trees (GBTs) and random forests (RFs) [65]. Xu et al. [66] used DL to establish a DILI prediction model with chemical structure data. In this study, a recursive neural network was used to build the model. The training was conducted on 475 drugs, and the external validation set of about 200 drugs was predicted. The accuracy, sensitivity, and specificity were 86.9%, 82.5%, and 92.9%, respectively. The performance of this model is better than the DILI prediction model previously reported. Lee et al. [67] employed machine learning models—including random forest (RF), logistic regression (LR), and Neural Networks (NNs)—to predict DILI, identify key molecular descriptors, and enhance model interpretability through feature importance ranking and LR coefficients.

In summary, drug hepatotoxicity prediction is primarily approached as a classification task, though data quality challenges complicate the identification of truly DILI-free drugs. Drug dosage, formulation, and patient baseline conditions can all influence prediction outcomes. Future research will focus on leveraging ML and DL for hepatotoxicity prediction through multi-source data integration, enhanced model interpretability, optimized deep learning architectures, multiclass and hierarchical prediction, personalized precision toxicology, and clinical/regulatory applications.

4.2.2. Nephrotoxicity

The kidney plays a key role in drug excretion, and nephrotoxicity is an important issue in drug research and development [68,69]. The mechanisms of drug-induced nephrotoxicity include the damage to renal tubular epithelial cells, the influence on glomerular filtration function, and the initiation of renal interstitial inflammation. In medical practice, the use of diuretics [70], antibiotics [71,72], and other drugs [73], as well as various imaging examinations, chemotherapy, radiotherapy, and interventional therapy, may lead to nephrotoxicity, which means that nephrotoxicity is difficult to capture. Therefore, we urgently need to build a model that can effectively predict drug-induced nephrotoxicity in advance, so as to achieve early intervention and nursing for patients. In the research on nephrotoxicity prediction, machine learning methods [74–76], such as random forest and logistic regression, combined with molecular descriptors and physical and chemical properties, are used in predictions. In recent years, deep learning models have also been gradually applied to the prediction of nephrotoxicity [77,78], such as the model based on convolutional neural networks (CNNs) [79], which extracts features for the prediction of nephrotoxicity via convolution operations on the two-dimensional structure image of drug molecules. At the same time, researchers have also tried to use metabonomic data [80] to analyze the relationship between small-molecule metabolites produced by drug metabolism in vivo and nephrotoxicity, providing new ideas for the prediction of nephrotoxicity. Banerjee et al. employed the Classification Read-Across Structure–Activity Relationship (c-RASAR) approach

to develop machine learning models using a recently curated dataset of orally active drugs' nephrotoxicity potential [76]. Their analysis revealed a strong overall performance for the c-RASAR models. The top-performing model, an LDA-based c-RASAR model utilizing topological descriptors, achieved Matthews Correlation Coefficient (MCC) values of 0.229 (training set) and 0.431 (test set). This model successfully predicted nephrotoxicity in a true external dataset derived from known nephrotoxic compounds in DrugBank DB, demonstrating good predictivity. Rao et al. collected a dataset of 360 FDA-classified compounds (231 non-nephrotoxic and 129 nephrotoxic) and predicted 6064 non-targeted interactions, of which 59% were validated in vitro. At the same time, they also calculated 55 physicochemical properties of these compounds. Multiple algorithms such as SVM, RF, and NN were used to establish ML models. They then combined these models into an integrated model to enhance performance [74]. The model they developed provides a promising early screening tool for identifying compounds with lower risk, thereby promoting safer drug development.

In summary, using clinical data and electronic health records, ML models can accurately predict acute kidney injury by learning and analyzing data patterns [81–83]. With the continuous development of new DL models, the integration of multimodal data [84] (such as images and physiological parameters) will be able to improve the performance of nephrotoxicity prediction.

4.2.3. Cardiotoxicity

Drug cardiotoxicity is one of the main reasons for drug development failure and delisting after-market launch. According to FDA statistics, about 30% of drugs are terminated or restricted in clinical trials or post-market stages due to cardiac safety issues, resulting in economic losses exceeding \$10 billion annually. Cardiac toxicity is mainly manifested as QT interval prolongation (caused by hERG channel inhibition) [85], cardiomyopathy, arrhythmia, etc. Its prediction and evaluation are the core challenges of drug development. AI-assisted drug toxicity discovery is an effective solution to reduce costs and facilitate the development of lead drug candidates. The most used cardiotoxicity datasets are Tox21, ChEMBL, ICE, PubChem, and hERGCentral. Yang et al. developed a large Language Model as Tools for Molecular Toxicity Prediction to predict cardiotoxicity [86]. They used GPT-4 combined with molecular docking technology to study the cardiotoxicity of three specific targets and examined Chinese medicinal materials listed as food and drugs. This study emphasizes the potential of ChatGPT in predicting molecular properties and its importance in medicinal chemistry, demonstrating that it promotes a new research paradigm: with datasets, high-precision learning models can be generated without computational knowledge or coding skills, making them easy to access and use. The human ether-a-go-go-related gene (hERG) potassium channel is pivotal in drug discovery due to its susceptibility to blockage by drug candidate molecules, which can cause severe cardiotoxic effects. Consequently, identifying and excluding potential hERG channel blockers at the earliest stages of drug development is crucial. To address the need for more precise, quantitative predictions, they developed hERGBoost [87], a cutting-edge machine learning model employing a gradient-boosting algorithm. This model demonstrates superior accuracy in predicting the IC₅₀ of drug candidates. Numerous studies are continuously innovating technology to more accurately predict cardiac toxicity [88–91]. AI is expected to improve the accuracy of preclinical prediction of cardiac toxicity to over 90%, significantly reducing drug development risks and patient safety threats.

Above all, cardiovascular toxicity is a well-known complication of thoracic radiotherapy (RT), which leads to an increase in incidence rate and mortality. However, existing

techniques for predicting cardiovascular toxicity have limitations. Predictive biomarkers [92] for cardiovascular toxicity may help maximize patient prognosis.

4.2.4. Neurotoxicity

Neurotoxicity involves the effects of drugs on the nervous system, which may lead to adverse consequences such as nerve damage and cognitive impairment. The core mechanisms of drug neurotoxicity (such as Parkinson's-like injury and Alzheimer's disease-related pathology) involve blood–brain barrier penetration, neurotransmitter interference (such as dopamine/glutamate imbalance), mitochondrial dysfunction, and microglial activation. Zhao et al. utilized real-world human clinical data to extract drug neurotoxicity information, developing 35 distinct classifiers by combining five machine learning methods with seven molecular fingerprinting techniques [93]. Among these, the MACCS-SVM model demonstrated optimal performance. Their analysis identified 18 structural alerts (SA) linked to neurotoxicity, offering interpretable insights. By leveraging both the three-dimensional structural information of drug molecules and their biological activity data on nerve cells, this approach enables the effective identification of compounds with potential neurotoxicity risks, thereby supporting the development and safety assessment of neuro pharmaceuticals. There are fewer neurotoxicity datasets, and commonly used datasets for neurotoxicity prediction are SIDER and PubChem. Chemical-induced neurotoxicity is a key aspect of chemical safety assessment. Traditional and expensive experimental methods require the development of high-throughput virtual screening. However, the small neurotoxic datasets limit the application of advanced deep learning techniques. Pang et al. developed a hybrid deep learning architecture called NeuTox 2.0 [94], which combines transfer learning based on self-supervised learning, graph neural networks, and molecular fingerprints/descriptors. Compared to traditional machine learning, NeuTox 2.0 has excellent noise resistance and high prediction accuracy. NeuTox 2.0 can be used as an effective tool for early neurotoxicity screening of environmental chemicals. HE et al. [95] combined various machine learning algorithms with molecular representations based on Particle Swarm Optimization-weighted scores to construct an ensemble prediction model called NeuTox. For the test set, NeuTox demonstrated excellent performance with an accuracy of 0.9064, outperforming the best-performing individual model. Identifying the structural features of chemical neurotoxicity facilitates the early design of non-toxic chemicals.

In conclusion, traditional compound toxicity detection techniques generally require the use of biochemical experiments, cell experiments, and even animal models, which not only consume a lot of time but also have high costs. The use of computational models for the toxicity prediction of organic compounds requires less input but yields enormous output. In particular, computational models based on the physical, chemical, and structural properties of compounds can even predict them before synthesis, greatly improving efficiency and making them increasingly popular.

4.2.5. Other Toxicities

Mutagenicity refers to the ability of chemical substances to induce gene mutations, which may lead to hereditary Deoxyribonucleic acid (DNA) damage and increase the risk of cancer. Standard assessment methods include the Ames test [96], which measures reverse mutations in strains (such as *Salmonella typhimurium*). The key databases for predicting mutability include OCHEM, which is a prediction based on structural alerts on the basis of QSAR. Rane et al. used the QSAR method to predict the mutagenicity of gliazoleozinc dimer impurities and simultaneously conducted in vitro evaluations through Ames and micronucleus tests [97]. The mutagenicity of dimer impurities in the in vitro method and

QSAR analysis was consistent with the computer prediction. Moreover, Carcinogenicity refers to the possibility of chemical substances inducing tumors. Trihalomethanes (THMs) in drinking water are regulated for carcinogenic health risks. However, frequent water quality monitoring imposes significant resource burdens. Li et al. proposed a framework [98] that combines interpretable machine learning with virtual data augmentation to predict the occurrence of THM and related cancer risks. These findings support data-driven soft sensing for water quality and health risk management in data-limited areas. Moreover, skin Corrosion/Irritation (Corr./Irrit.) has long been a health hazard in the Globally Harmonized System (GHS). Huang et al. have established several computer models to predict skin allergies as an alternative to increasingly limited animal experiments [99]. We believe that the establishment of databases and computer models can facilitate the assessment of chemical safety and related research. Serious eye damage and eye irritation have been authenticated to be significant human health issues in various fields such as ophthalmic pharmaceuticals. Due to the shortcomings of traditional animal testing methods, *in silico* methods have advanced to study eye toxicity. Di et al. developed models for predicting severe eye damage and eye irritation potential of compounds using 2299 and 5214 compounds, respectively. The test accuracy of the best serious eye damage model and eye irritation model reached 0.972 and 0.959, respectively [100]. In total, the prediction models and structural alerts contributed to providing hazard identification and assessing chemical safety.

5. Challenges and Prospects

While AI-driven toxicity prediction offers transformative potential, critical limitations in data quality, model design, and translational validation hinder its reliability in real-world drug development. These challenges demand urgent interdisciplinary collaboration.

5.1. Expanding Sample Sizes

Insufficient sample size constitutes a primary challenge, compromising model training and adversely impacting prediction accuracy and generalization capability. As noted, toxicity endpoint labels (e.g., clinical toxicity, nephrotoxicity) remain scarce. However, advances in sequencing technology are rapidly generating multi-omics data and drug perturbation-based toxicity datasets [101]. Machine learning techniques—including data augmentation, pretrained models, active learning, and meta-learning—can mitigate data limitations [102]. Furthermore, data sharing and integration initiatives substantially expand sample sizes. Platforms like CTD and TTD aggregate multi-source data, while public participation and open-access policies accelerate dataset growth. Collectively, these strategies enhance the performance and reliability of ML in toxicity prediction.

5.2. Improving Data Quality and Integrating Diverse Sources

The scarcity of high-quality data and the challenges of integrating multi-source datasets have hindered the broader application of machine learning. To address this, data from diverse sources must be standardized using relatively uniform approaches to ensure interoperability. Implementing comprehensive, rigorous data-cleaning and cross-validation processes is essential to minimize noise and errors. Integrating heterogeneous data sources for toxicity prediction faces significant hurdles, including data quality issues, heterogeneity, and sparsity. Ensemble learning methods, along with techniques such as federated learning (FL) and multimodal ML, facilitate the integration of toxicological, chemical, and multi-omics data, helping to mitigate data scarcity [103]. Synthetic data generation technology creates artificial data that simulate the features of unknown categories. Models trained on this synthetic data, which share statistical attributes with real-world data, show improved

performance on actual datasets. This technology provides an effective solution to data scarcity and enhances the model's capability to identify unseen categories. Personalized toxicity prediction, leveraging patient-specific data or synthetic data generation, further addresses data shortages. Collectively, these methods enable the effective combination and enhancement of multi-source datasets, creating unified, high-quality data resources.

5.3. Enhancing ML Models for Diverse Toxicity Detection

Predicting multiple toxicities (such as acute toxicity, carcinogenicity, and organ-specific toxicity) is challenging due to their respective unique biological mechanisms, which require the use of specialized models. The prediction of different toxicity endpoints usually requires different feature sets. For example, chemical toxicity might be more suitable to be characterized by molecular descriptors [104]. However, biological toxicity depends more on the gene expression profile [105]. Take liver toxicity as an example. Its prediction requires the integration of liver-specific multi-omics data such as metabolic enzyme activities, bile acid circulation, and oxidative stress and the combination of dynamic response models. Therefore, to further improve the prediction performance, it is necessary to construct appropriate feature representations and dedicated models for various toxicities. DL models offer new opportunities to address the challenges of various toxicity predictions. Compared with traditional methods, DL technology can capture complex nonlinear relationships in data more effectively and has excellent scalability, capable of handling large datasets—which is particularly important for tasks involving high-dimensional chemical and biological data. The key to developing a DL workflow that integrates patients' genetic information and clinical data to achieve personalized prediction and optimize treatment outcomes lies in promoting interdisciplinary collaboration among toxicologists, data scientists, biologists, and clinicians. More benchmark databases are urgently needed to construct and validate toxicity prediction models. However, existing databases often hinder repeatability and model comparison due to the lack of standardization. Shared and well-planned benchmarks (such as TOXRIC, TDC, and MoleculeNet) can enhance the consistency of evaluations. We encourage researchers to enhance repeatability through open-source code, hyperparameters, and data preprocessing steps on platforms such as GitHub and Zenodo. Ultimately, formulating standardized benchmarks and reporting guidelines is crucial for ensuring the reliability of results and accelerating progress in the field.

5.4. Enhance the Interpretability of ML in the Prediction of Drug Toxicity

In drug toxicity prediction, a major challenge for ML models is their lack of interpretability. The core objective of explainability is to present a transparent and traceable prediction process, enabling researchers to understand model decision-making. This enhances the credibility, usability, and acceptability of models in clinical decision-making and regulatory approval, which is crucial for related work.

While advanced interpretability methods like Shapley additive explanations (SHAPs) [106] can handle multiple data types, they often require customization when applied to complex biological data (e.g., time series, multimodal, and high-dimensional sparse data). Furthermore, the absence of widely accepted evaluation criteria makes it difficult to quantify or assess these methods effectively. Current evaluations are largely limited to simple qualitative analyses or narrow-scoped quantitative metrics, falling short of practical application demands [107]. The inherent diversity and complexity of biomedical data (genomic, omics, and clinical) further compound these challenges.

As artificial intelligence overcomes these hurdles and aligns with regulatory frameworks, its potential in regulatory toxicology is being realized. For ML models to meet regulatory requirements, they must ensure transparency and repeatability, gaining ap-

proval through standardized processes. AI promotes the development of reliable in vitro and computational models, potentially reducing reliance on animal experimentation. These advancements have cemented machine learning's potential as a cornerstone of modern regulatory toxicology practice [108]. Integrating multi-omics and chemical data improves toxicity prediction accuracy and toxicological mechanism understanding. Automated, interpretable models can assist regulatory agencies in prioritizing chemicals based on quantitative predictions and guiding subsequent testing. These advancements have cemented machine learning's role as a cornerstone of modern regulatory toxicology practice [109].

6. Conclusions

In this review, we emphasize the increasingly critical role of AI in addressing challenges associated with drug toxicity prediction. By classifying diverse drug-induced toxicities and analyzing their characteristics alongside predictive models, we demonstrate how ML bridges the gap between predictive accuracy and mechanistic understanding. Furthermore, we comprehensively summarize key databases and tools relevant to drug toxicity, providing researchers with valuable resources to explore toxicity across various domains and facilitating data-driven insights for informed decision-making. These advances highlight ML's potential to overcome limitations inherent in traditional toxicity assessment methods—such as high costs, ethical concerns, and cross-species variability. Despite this progress, significant challenges persist, including data scarcity, model interpretability, and multi-source data integration. Addressing these issues necessitates the continual development of robust, interpretable ML models underpinned by standardized evaluation metrics.

Author Contributions: X.Z. and B.L. contributed to conception and design of the study. H.W., Z.L., and R.Z. conducted a literature survey. R.Z. interpreted the results and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors gratefully acknowledge the encouragement and support from Chaoyue Yang. Some cartoon components were from www.figdraw.com, accessed on 19 April 2025.

Conflicts of Interest: The authors declare that this study does not have any potential conflicts of interest.

List of Abbreviations

The list of abbreviations is shown as follows.

Full Term	Abbreviation
Absorption, Distribution, Metabolism, Excretion and Toxicity	ADMET
Application Programming Interface	API
Artificial Intelligence	AI
Classification Read-Across Structure-Activity Relationship	RASAR
Comparative Toxicogenomics Database	CTD
Convolutional Neural Network	CNN
Deep Learning	DL
Deep Neural Network	DNN

Deoxyribonucleic Acid	DNA
Distributed Structure-Searchable Toxicity Database	DSSTox
Drug-Induced Liver Injury	DILI
Electronic Medical Record	EMR
Environmental Protection Agency	EPA
European Bioinformatics Institute	EBI
Extreme Gradient Boosting	XGB
FDA Adverse Event Reporting System	FAERS
Federated Learning	FL
Food and Drug Administration	FDA
Gene Expression Nebulas	GEN
Gene Expression Omnibus	GEO
Globally Harmonized System	GHS
Gradient Boosting Machines	GBM
Gradient Boosting Trees	GBT
Graph Convolutional Networks	GCN
Graph Neural Network	GNN
Half Maximal Inhibitory Concentration	IC50
Integrated Chemical Environment	ICE
K-Nearest Neighbor	KNN
Kyoto Encyclopedia of Genes and Genomes	KEGG
LiverTox: Clinical and Research Information on Drug-Induced Liver Injury	LIVERTOX
Logistic Regression	LR
Machine Learning	ML
Matthews Correlation Coefficient	MCC
Median Lethal Dose	LD50
Molecular Access System	MACCS
National Cancer Institute	NCI
National Institute of Technology and Evaluation	NITE
National Library of Medicine	NIH
National Toxicology Program	NTP
Natural Language Processing	NLP
Neural Network	NN
Online Chemical Modeling Environment	OCHEM
Quantitative Structure-Activity Relationship	QSAR
Radio Therapy	RT
Random Forest	RF
Root Mean Square Errors	RMSEs
SHapley Additive exPlanations	SHAP
Side Effect Resource	SIDER
Structural Alarms	SA
Support Vector Machine	SVM
The Cancer Genome Atlas	TCGA
Therapeutic Target Database	TTD
Therapeutics Data Commons	TDC
Toxicity Reference Database	ToxRefDB
Toxicology Resources for Intelligent Computation	TOXRIC
Toxin and Toxin Target Database	T3DB
Trihalomethanes	THMs
Universal Protein Database	UniProt
World Health Organization	WHO

References

- Basile, A.O.; Yahi, A.; Tatonetti, N.P. Artificial Intelligence for Drug Toxicity and Safety. *Trends Pharmacol. Sci.* **2019**, *40*, 624–635. [\[CrossRef\]](#) [\[PubMed\]](#)
- Raies, A.B.; Bajic, V.B. In silico toxicology: Computational methods for the prediction of chemical toxicity. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2016**, *6*, 147–172. [\[CrossRef\]](#) [\[PubMed\]](#)
- Onakpoya, I.J.; Heneghan, C.J.; Aronson, J.K. Worldwide withdrawal of medicinal products because of adverse drug reactions: A systematic review and analysis. *Crit. Rev. Toxicol.* **2016**, *46*, 477–489. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lin, N.; Zhou, X.; Geng, X.; Drewell, C.; Hubner, J.; Li, Z.; Zhang, Y.; Xue, M.; Marx, U.; Li, B. Repeated dose multi-drug testing using a microfluidic chip-based coculture of human liver and kidney proximal tubules equivalents. *Sci. Rep.* **2020**, *10*, 8879. [\[CrossRef\]](#)
- Handelman, G.S.; Kok, H.K.; Chandra, R.V.; Razavi, A.H.; Lee, M.J.; Asadi, H. eDoctor: Machine learning and the future of medicine. *J. Intern. Med.* **2018**, *284*, 603–619. [\[CrossRef\]](#)
- Ali, H.; Shroff, A.; Fulop, T.; Molnar, M.Z.; Sharif, A.; Burke, B.; Shroff, S.; Briggs, D.; Krishnan, N. Artificial intelligence assisted risk prediction in organ transplantation: A UK Live-Donor Kidney Transplant Outcome Prediction tool. *Ren. Fail.* **2025**, *47*, 2431147. [\[CrossRef\]](#)
- Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J.T. Deep learning for healthcare: Review, opportunities and challenges. *Brief. Bioinform.* **2018**, *19*, 1236–1246. [\[CrossRef\]](#)
- Mullowney, M.W.; Duncan, K.R.; Elsayed, S.S.; Garg, N.; van der Hooft, J.J.J.; Martin, N.I.; Meijer, D.; Terlouw, B.R.; Biermann, F.; Blin, K.; et al. Artificial intelligence for natural product drug discovery. *Nat. Rev. Drug Discov.* **2023**, *22*, 895–916. [\[CrossRef\]](#)
- Wu, L.; Yan, B.; Han, J.; Li, R.; Xiao, J.; He, S.; Bo, X. TOXRIC: A comprehensive database of toxicological data and benchmarks. *Nucleic Acids Res.* **2023**, *51*, D1432–D1445. [\[CrossRef\]](#)
- Bell, S.M.; Phillips, J.; Sedykh, A.; Tandon, A.; Sprankle, C.; Morefield, S.Q.; Shapiro, A.; Allen, D.; Shah, R.; Maull, E.A.; et al. An Integrated Chemical Environment to Support 21st-Century Toxicology. *Environ. Health Perspect.* **2017**, *125*, 054501. [\[CrossRef\]](#)
- Daniel, A.B.; Choksi, N.; Abedini, J.; Bell, S.; Ceger, P.; Cook, B.; Karmaus, A.L.; Rooney, J.; To, K.T.; Allen, D.; et al. Data curation to support toxicity assessments using the Integrated Chemical Environment. *Front. Toxicol.* **2022**, *4*, 987848. [\[CrossRef\]](#)
- Grulke, C.M.; Williams, A.J.; Thillanadarajah, I.; Richard, A.M. EPA's DSSTox database: History of development of a curated chemistry resource supporting computational toxicology research. *Comput. Toxicol.* **2019**, *12*, 100096. [\[CrossRef\]](#) [\[PubMed\]](#)
- Knox, C.; Wilson, M.; Klinger, C.M.; Franklin, M.; Oler, E.; Wilson, A.; Pon, A.; Cox, J.; Chin, N.E.L.; Strawbridge, S.A.; et al. DrugBank 6.0: The DrugBank Knowledgebase for 2024. *Nucleic Acids Res.* **2024**, *52*, D1265–D1275. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zdrzil, B.; Felix, E.; Hunter, F.; Manners, E.J.; Blackshaw, J.; Corbett, S.; de Veij, M.; Ioannidis, H.; Lopez, D.M.; Mosquera, J.F.; et al. The ChEMBL Database in 2023: A drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res.* **2024**, *52*, D1180–D1192. [\[CrossRef\]](#)
- Kovalishyn, V.; Zyabrev, V.; Kachaeva, M.; Ziabrev, K.; Keith, K.; Harden, E.; Hartline, C.; James, S.H.; Brovarets, V. Design of new imidazole derivatives with anti-HCMV activity: QSAR modeling, synthesis and biological testing. *J. Comput. Aided Mol. Des.* **2021**, *35*, 1177–1187. [\[CrossRef\]](#)
- Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B.; et al. PubChem 2023 update. *Nucleic Acids Res.* **2023**, *51*, D1373–D1380. [\[CrossRef\]](#)
- Ito, M.; Codony-Servat, C.; Codony-Servat, J.; Llige, D.; Chaib, I.; Sun, X.; Miao, J.; Sun, R.; Cai, X.; Verlicchi, A.; et al. Targeting PKC α -PAK1 signaling pathways in EGFR and KRAS mutant adenocarcinoma and lung squamous cell carcinoma. *Cell Commun. Signal* **2019**, *17*, 137. [\[CrossRef\]](#)
- Pu, X.; Ye, Q.; Cai, J.; Yang, X.; Fu, Y.; Fan, X.; Wu, H.; Chen, J.; Qiu, Y.; Yue, S. Typing FGFR2 translocation determines the response to targeted therapy of intrahepatic cholangiocarcinomas. *Cell Death Dis.* **2021**, *12*, 256. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liu, H.; Yang, Q.; Li, Z.; Yan, S.; Ming, S. Systematic analysis of sugammadex-related adverse drug reaction signals using FAERS database. *Int. J. Surg.* **2025**, *111*, 1988–1994. [\[CrossRef\]](#)
- Elizabeth, R.M.C.; Sattari, F.; Lefsrud, L.; Gue, B. Visualizing what's missing: Using deep learning and Bow-Tie diagrams to identify and visualize missing leading indicators in industrial construction. *J. Saf. Res.* **2025**, *93*, 1–11. [\[CrossRef\]](#)
- Feshuk, M.; Kolaczowski, L.; Dunham, K.; Davidson-Fritz, S.E.; Carstens, K.E.; Brown, J.; Judson, R.S.; Paul Friedman, K. The ToxCast pipeline: Updates to curve-fitting approaches and database structure. *Front. Toxicol.* **2023**, *5*, 1275980. [\[CrossRef\]](#) [\[PubMed\]](#)
- Williams, A.J.; Grulke, C.M.; Edwards, J.; McEachran, A.D.; Mansouri, K.; Baker, N.C.; Patlewicz, G.; Shah, I.; Wambaugh, J.F.; Judson, R.S.; et al. The CompTox Chemistry Dashboard: A community data resource for environmental chemistry. *J. Cheminform.* **2017**, *9*, 61. [\[CrossRef\]](#)
- Feshuk, M.; Kolaczowski, L.; Watford, S.; Paul Friedman, K. ToxRefDB v2.1: Update to curated in vivo study data in the Toxicity Reference Database. *Front. Toxicol.* **2023**, *5*, 1260305. [\[CrossRef\]](#)

24. Wishart, D.; Arndt, D.; Pon, A.; Sajed, T.; Guo, A.C.; Djoumbou, Y.; Knox, C.; Wilson, M.; Liang, Y.; Grant, J.; et al. T3DB: The toxic exposome database. *Nucleic Acids Res.* **2015**, *43*, D928–D934. [\[CrossRef\]](#)
25. Zhou, Y.; Zhang, Y.; Zhao, D.; Yu, X.; Shen, X.; Zhou, Y.; Wang, S.; Qiu, Y.; Chen, Y.; Zhu, F. TTD: Therapeutic Target Database describing target druggability information. *Nucleic Acids Res.* **2024**, *52*, D1465–D1477. [\[CrossRef\]](#)
26. Kuhn, M.; Letunic, I.; Jensen, L.J.; Bork, P. The SIDER database of drugs and side effects. *Nucleic Acids Res.* **2016**, *44*, D1075–D1079. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Zhang, Y.; Zhou, Y.; Zhou, Y.; Yu, X.; Shen, X.; Hong, Y.; Zhang, Y.; Wang, S.; Mou, M.; Zhang, J.; et al. TheMarker: A comprehensive database of therapeutic biomarkers. *Nucleic Acids Res.* **2024**, *52*, D1450–D1464. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Davis, A.P.; Wiegiers, T.C.; Johnson, R.J.; Sciaky, D.; Wiegiers, J.; Mattingly, C.J. Comparative Toxicogenomics Database (CTD): Update 2023. *Nucleic Acids Res.* **2023**, *51*, D1257–D1262. [\[CrossRef\]](#)
29. Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K.H.; Sherman, P.M.; Holko, M.; et al. NCBI GEO: Archive for functional genomics data sets--update. *Nucleic Acids Res.* **2013**, *41*, D991–D995. [\[CrossRef\]](#)
30. Lin, L.; Wan, L.; He, H.; Liu, W. Drug vector representation: A tool for drug similarity analysis. *Mol. Genet. Genom.* **2020**, *295*, 1055–1062. [\[CrossRef\]](#)
31. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361. [\[CrossRef\]](#) [\[PubMed\]](#)
32. UniProt, C. UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489. [\[CrossRef\]](#)
33. Huang, K.; Fu, T.; Gao, W.; Zhao, Y.; Roohani, Y.; Leskovec, J.; Coley, C.W.; Xiao, C.; Sun, J.; Zitnik, M. Artificial intelligence foundation for therapeutic science. *Nat. Chem. Biol.* **2022**, *18*, 1033–1036. [\[CrossRef\]](#)
34. Kohonen, P.; Benfenati, E.; Bower, D.; Ceder, R.; Crump, M.; Cross, K.; Grafstrom, R.C.; Healy, L.; Helma, C.; Jeliaskova, N.; et al. The ToxBank Data Warehouse: Supporting the Replacement of In Vivo Repeated Dose Systemic Toxicity Testing. *Mol. Inform.* **2013**, *32*, 47–63. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Kelleci Celi, K.F.; Karaduman, G. Machine Learning-Based Prediction of Drug-Induced Hepatotoxicity: An OvA-QSTR Approach. *J. Chem. Inf. Model.* **2023**, *63*, 4602–4614. [\[CrossRef\]](#)
36. Zhang, Y.; Zou, D.; Zhu, T.; Xu, T.; Chen, M.; Niu, G.; Zong, W.; Pan, R.; Jing, W.; Sang, J.; et al. Gene Expression Nebulas (GEN): A comprehensive data portal integrating transcriptomic profiles across multiple species at both bulk and single-cell levels. *Nucleic Acids Res.* **2022**, *50*, D1016–D1024. [\[CrossRef\]](#)
37. Tomczak, K.; Czerwinska, P.; Wiznerowicz, M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp. Oncol.* **2015**, *19*, A68–A77. [\[CrossRef\]](#)
38. Wu, X.; Wei, D.; Zhou, Y.; Cao, Q.; Han, G.; Han, E.; Chen, Z.; Guo, Y.; Huo, W.; Wang, C.; et al. Pesticide exposures and 10-year atherosclerotic cardiovascular disease risk: Integrated epidemiological and bioinformatics analysis. *J. Hazard. Mater.* **2025**, *485*, 136835. [\[CrossRef\]](#)
39. Li, X.; Cui, Y.; Gao, S.; Zhang, Q.; Dai, Y.; Wang, S.; Wu, J.; Li, G.; Song, J. Development and validation of a score model for predicting the risk of first esophagogastric variceal hemorrhage and mortality in patients with hepatocellular carcinoma. *Ann. Med.* **2025**, *57*, 2490210. [\[CrossRef\]](#)
40. Albrich, W.C.; Just, N.; Kahlert, C.; Casanova, C.; Baty, F.; Hilty, M. Serotype epidemiology and case-fatality risk of invasive pneumococcal disease: A nationwide population study from Switzerland, 2012–2022. *Emerg. Microbes Infect.* **2025**, *14*, 2488189. [\[CrossRef\]](#)
41. Ai, S.; Liu, X.; Zhao, N.; Miao, Q.; Qin, Y.; Li, X. Risk factors of infective endocarditis-associated acute kidney injury: Benefits of low-dose amikacin and surgery. *Ann. Med.* **2025**, *57*, 2482023. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Luo, Y.; Ding, W.; Yang, X.; Bai, H.; Jiao, F.; Guo, Y.; Zhang, T.; Zou, X.; Wang, Y. Construction and validation of a predictive model for meningoenephalitis in pediatric scrub typhus based on machine learning algorithms. *Emerg. Microbes Infect.* **2025**, *14*, 2469651. [\[CrossRef\]](#)
43. Hwang, J.; Youm, C.; Park, H.; Kim, B.; Choi, H.; Cheon, S.M. Machine learning for early detection and severity classification in people with Parkinson's disease. *Sci. Rep.* **2025**, *15*, 234. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Cakir, A.; Tuncer, M.; Taymaz-Nikerel, H.; Ulucan, O. Side effect prediction based on drug-induced gene expression profiles and random forest with iterative feature selection. *Pharmacogenom. J.* **2021**, *21*, 673–681. [\[CrossRef\]](#)
45. Schoning, V.; Krahenbuhl, S.; Drewe, J. The hepatotoxic potential of protein kinase inhibitors predicted with Random Forest and Artificial Neural Networks. *Toxicol. Lett.* **2018**, *299*, 145–148. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Lesinski, W.; Mnich, K.; Golinska, A.K.; Rudnicki, W.R. Integration of human cell lines gene expression and chemical properties of drugs for Drug Induced Liver Injury prediction. *Biol. Direct* **2021**, *16*, 2. [\[CrossRef\]](#)
47. Fourches, D.; Barnes, J.C.; Day, N.C.; Bradley, P.; Reed, J.Z.; Tropsha, A. Cheminformatics analysis of assertions mined from literature that describe drug-induced liver injury in different species. *Chem. Res. Toxicol.* **2010**, *23*, 171–183. [\[CrossRef\]](#) [\[PubMed\]](#)

48. Mulliner, D.; Schmidt, F.; Stolte, M.; Spirkel, H.P.; Czich, A.; Amberg, A. Computational Models for Human and Animal Hepatotoxicity with a Global Application Scope. *Chem. Res. Toxicol.* **2016**, *29*, 757–767. [\[CrossRef\]](#)
49. Lee, S.; Kang, Y.M.; Park, H.; Dong, M.S.; Shin, J.M.; No, K.T. Human nephrotoxicity prediction models for three types of kidney injury based on data sets of pharmacological compounds and their metabolites. *Chem. Res. Toxicol.* **2013**, *26*, 1652–1659. [\[CrossRef\]](#)
50. Cai, C.; Guo, P.; Zhou, Y.; Zhou, J.; Wang, Q.; Zhang, F.; Fang, J.; Cheng, F. Deep Learning-Based Prediction of Drug-Induced Cardiotoxicity. *J. Chem. Inf. Model.* **2019**, *59*, 1073–1084. [\[CrossRef\]](#)
51. Weber, S.; Erhardt, F.; Allgeier, J.; Saka, D.; Donga, N.; Neumann, J.; Lange, C.M.; Gerbes, A.L. Drug-Induced Liver Injury Caused by Metamizole: Identification of a Characteristic Injury Pattern. *Liver Int.* **2025**, *45*, e70012. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Zhang, H.; Yu, P.; Zhang, T.G.; Kang, Y.L.; Zhao, X.; Li, Y.Y.; He, J.H.; Zhang, J. In silico prediction of drug-induced myelotoxicity by using Naive Bayes method. *Mol. Divers.* **2015**, *19*, 945–953. [\[CrossRef\]](#)
53. Kokori, E.; Patel, R.; Olatunji, G.; Ukoaka, B.M.; Abraham, I.C.; Ajekigbe, V.O.; Kwape, J.M.; Babalola, A.E.; Udam, N.G.; Aderinto, N. Machine learning in predicting heart failure survival: A review of current models and future prospects. *Heart Fail. Rev.* **2025**, *30*, 431–442. [\[CrossRef\]](#)
54. Vo, A.H.; Van Vleet, T.R.; Gupta, R.R.; Liguori, M.J.; Rao, M.S. An Overview of Machine Learning and Big Data for Drug Toxicity Evaluation. *Chem. Res. Toxicol.* **2020**, *33*, 20–37. [\[CrossRef\]](#)
55. Raies, A.B.; Bajic, V.B. In silico toxicology: Comprehensive benchmarking of multi-label classification methods applied to chemical toxicity data. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1352. [\[CrossRef\]](#)
56. Zhang, W.; Liu, F.; Luo, L.; Zhang, J. Predicting drug side effects by multi-label learning and ensemble learning. *BMC Bioinform.* **2015**, *16*, 365. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Kim, J.; Shin, M. An integrative model of multi-organ drug-induced toxicity prediction using gene-expression data. *BMC Bioinform.* **2014**, *15* (Suppl 16), S2. [\[CrossRef\]](#)
58. Zhong, S.; Guan, X. Count-Based Morgan Fingerprint: A More Efficient and Interpretable Molecular Representation in Developing Machine Learning-Based Predictive Regression Models for Water Contaminants' Activities and Properties. *Environ. Sci. Technol.* **2023**, *57*, 18193–18202. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Ali, S.M.; Atta, S.; Naz, I.; Fatima, H.; Haq, I.U. Toxicological assessment of standardized *Artemisia roxburghiana* wall. ex Besser aqueous leaf extract to acute and subacute exposure in Albino mice. *J. Ethnopharmacol.* **2025**, *344*, 119531. [\[CrossRef\]](#)
60. Lou, S.; Yu, Z.; Huang, Z.; Wang, H.; Pan, F.; Li, W.; Liu, G.; Tang, Y. In Silico Prediction of Chemical Acute Dermal Toxicity Using Explainable Machine Learning Methods. *Chem. Res. Toxicol.* **2024**, *37*, 513–524. [\[CrossRef\]](#)
61. Li, Y.; Wang, B.; Ma, F.; Fan, W.; Wang, Y.; Chen, L.; Dong, Z. Using the super-learner to predict the chemical acute toxicity on rats. *J. Hazard. Mater.* **2024**, *480*, 136311. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Zeng, L.; Zhang, Y.; Mai, X.; Ai, P.; Xie, L.; Xue, B.; Zheng, Q. Excellent absorption-dominant electromagnetic interference shielding performances of asymmetric gradient layered composite films exploited with assistance of machine learning. *J. Colloid. Interface Sci.* **2025**, *697*, 137927. [\[CrossRef\]](#)
63. Feitosa, F.L.; Victoria, F.C.; Sanches, I.H.; Silva-Mendonca, S.; Borba, J.; Braga, R.C.; Andrade, C.H. Cyto-Safe: A Machine Learning Tool for Early Identification of Cytotoxic Compounds in Drug Discovery. *J. Chem. Inf. Model.* **2024**, *64*, 9056–9062. [\[CrossRef\]](#) [\[PubMed\]](#)
64. Chen, Z.; Jiang, Y.; Zhang, X.; Zheng, R.; Qiu, R.; Sun, Y.; Zhao, C.; Shang, H. ResNet18DNN: Prediction approach of drug-induced liver injury by deep neural network with ResNet18. *Brief. Bioinform.* **2022**, *23*, bbab503. [\[CrossRef\]](#)
65. Fuzi, B.; Mathai, N.; Kirchmair, J.; Ecker, G.F. Toxicity prediction using target, interactome, and pathway profiles as descriptors. *Toxicol. Lett.* **2023**, *381*, 20–26. [\[CrossRef\]](#)
66. Xu, Y.; Dai, Z.; Chen, F.; Gao, S.; Pei, J.; Lai, L. Deep Learning for Drug-Induced Liver Injury. *J. Chem. Inf. Model.* **2015**, *55*, 2085–2093. [\[CrossRef\]](#)
67. Lee, S.; Yoo, S. InterDILI: Interpretable prediction of drug-induced liver injury through permutation feature importance and attention mechanism. *J. Cheminform.* **2024**, *16*, 1. [\[CrossRef\]](#)
68. Gadaleta, D.; Garcia de Lomana, M.; Serrano-Candelas, E.; Ortega-Vallbona, R.; Gozalbes, R.; Roncaglioni, A.; Benfenati, E. Quantitative structure-activity relationships of chemical bioactivity toward proteins associated with molecular initiating events of organ-specific toxicity. *J. Cheminform.* **2024**, *16*, 122. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Jochum, K.; Miccoli, A.; Sommersdorf, C.; Poetz, O.; Braeuning, A.; Tralau, T.; Marx-Stoelting, P. Comparative case study on NAMs: Towards enhancing specific target organ toxicity analysis. *Arch. Toxicol.* **2024**, *98*, 3641–3658. [\[CrossRef\]](#)
70. Zhang, R.; Liu, Y.; Cao, J.; Lao, J.; Wang, B.; Li, S.; Huang, X.; Tang, F.; Li, X. The incidence and risk factors analysis of acute kidney injury in hospitalized patients received diuretics: A single-center retrospective study. *Front. Pharmacol.* **2022**, *13*, 924173. [\[CrossRef\]](#)
71. Song, M.H.; Xiang, B.X.; Yang, C.Y.; Lee, C.H.; Yan, Y.X.; Yang, Q.J.; Yin, W.J.; Zhou, Y.; Zuo, X.C.; Xie, Y.L. A pilot clinical risk model to predict polymyxin-induced nephrotoxicity: A real-world, retrospective cohort study. *J. Antimicrob. Chemother.* **2024**, *79*, 1919–1928. [\[CrossRef\]](#) [\[PubMed\]](#)

72. Bao, P.; Sun, Y.; Qiu, P.; Li, X. Development and validation of a nomogram to predict the risk of vancomycin-related acute kidney injury in critical care patients. *Front. Pharmacol.* **2024**, *15*, 1389140. [[CrossRef](#)]
73. Wang, Z.; Wang, Q.; Gong, X. Unveiling the Mysteries of Contrast-Induced Acute Kidney Injury: New Horizons in Pathogenesis and Prevention. *Toxics* **2024**, *12*, 620. [[CrossRef](#)]
74. Rao, M.; Nassiri, V.; Srivastava, S.; Yang, A.; Brar, S.; McDuffie, E.; Sachs, C. Artificial Intelligence and Machine Learning Models for Predicting Drug-Induced Kidney Injury in Small Molecules. *Pharmaceuticals* **2024**, *17*, 1550. [[CrossRef](#)] [[PubMed](#)]
75. Fu, L.; Zheng, X.; Luo, J.; Zhang, Y.; Gao, X.; Jin, L.; Liu, W.; Zhang, C.; Gao, D.; Xu, B.; et al. Machine learning accelerates the discovery of epitope-based dual-bioactive peptides against skin infections. *Int. J. Antimicrob. Agents* **2024**, *64*, 107371. [[CrossRef](#)]
76. Banerjee, A.; Roy, K. Machine learning assisted classification RASAR modeling for the nephrotoxicity potential of a curated set of orally active drugs. *Sci. Rep.* **2025**, *15*, 808. [[CrossRef](#)]
77. Heo, S.; Kang, E.A.; Yu, J.Y.; Kim, H.R.; Lee, S.; Kim, K.; Hwangbo, Y.; Park, R.W.; Shin, H.; Ryu, K.; et al. Time Series AI Model for Acute Kidney Injury Detection Based on a Multicenter Distributed Research Network: Development and Verification Study. *JMIR Med. Inform.* **2024**, *12*, e47693. [[CrossRef](#)]
78. Feng, X.; Jin, X.; Zhou, R.; Jiang, Q.; Wang, Y.; Zhang, X.; Shang, K.; Zhang, J.; Yu, C.; Shou, J. Deep learning approach identified a gene signature predictive of the severity of renal damage caused by chronic cadmium accumulation. *J. Hazard. Mater.* **2022**, *433*, 128795. [[CrossRef](#)] [[PubMed](#)]
79. Shi, Y.; Hua, Y.; Wang, B.; Zhang, R.; Li, X. In Silico Prediction and Insights Into the Structural Basis of Drug Induced Nephrotoxicity. *Front. Pharmacol.* **2021**, *12*, 793332. [[CrossRef](#)]
80. Devalckeneer, A.; Bouviez, M.; Gautier, A.; Colet, J.M. Metabolomic Prediction of Cadmium Nephrotoxicity in the Snail *Helix aspersa maxima*. *Metabolites* **2024**, *14*, 455. [[CrossRef](#)]
81. Mizuno, S.; Noda, T.; Mogushi, K.; Hase, T.; Iida, Y.; Takeuchi, K.; Ishiwata, Y.; Uchida, S.; Nagata, M. Prediction of Vancomycin-Associated Nephrotoxicity Based on the Area under the Concentration-Time Curve of Vancomycin: A Machine Learning Analysis. *Biol. Pharm. Bull.* **2024**, *47*, 1946–1952. [[CrossRef](#)] [[PubMed](#)]
82. Dzidic-Krivic, A.; Sher, E.K.; Kusturica, J.; Farhat, E.K.; Nawaz, A.; Sher, F. Unveiling drug induced nephrotoxicity using novel biomarkers and cutting-edge preventive strategies. *Chem. Biol. Interact.* **2024**, *388*, 110838. [[CrossRef](#)] [[PubMed](#)]
83. Chiu, L.W.; Ku, Y.E.; Chan, F.Y.; Lie, W.N.; Chao, H.J.; Wang, S.Y.; Shen, W.C.; Chen, H.Y. Machine learning algorithms to predict colistin-induced nephrotoxicity from electronic health records in patients with multidrug-resistant Gram-negative infection. *Int. J. Antimicrob. Agents* **2024**, *64*, 107175. [[CrossRef](#)]
84. Su, R.; Yang, H.; Wei, L.; Chen, S.; Zou, Q. A multi-label learning model for predicting drug-induced pathology in multi-organ based on toxicogenomics data. *PLoS Comput. Biol.* **2022**, *18*, e1010402. [[CrossRef](#)]
85. Mukherjee, S.; Swanson, K.; Walther, P.; Shivnaraine, R.V.; Leitz, J.; Pang, P.D.; Zou, J.; Wu, J.C. ADMET-AI Enables Interpretable Predictions of Drug-Induced Cardiotoxicity. *Circulation* **2025**, *151*, 285–287. [[CrossRef](#)]
86. Yang, H.; Xiu, J.; Yan, W.; Liu, K.; Cui, H.; Wang, Z.; He, Q.; Gao, Y.; Han, W. Large Language Models as Tools for Molecular Toxicity Prediction: AI Insights into Cardiotoxicity. *J. Chem. Inf. Model.* **2025**, *65*, 2268–2282. [[CrossRef](#)] [[PubMed](#)]
87. Yu, M.S.; Lee, J.; Lee, Y.; Cho, D.; Oh, K.S.; Jang, J.; Nong, N.T.; Lee, H.M.; Na, D. hERGBBoost: A gradient boosting model for quantitative IC(50) prediction of hERG channel blockers. *Comput. Biol. Med.* **2025**, *184*, 109416. [[CrossRef](#)]
88. Chen, Z.; Li, N.; Zhang, P.; Li, Y.; Li, X. CardioDPi: An explainable deep-learning model for identifying cardiotoxic chemicals targeting hERG, Cav1.2, and Nav1.5 channels. *J. Hazard. Mater.* **2024**, *474*, 134724. [[CrossRef](#)]
89. Au Yeung, V.P.W.; Obrezanova, O.; Zhou, J.; Yang, H.; Bowen, T.J.; Ivanov, D.; Saffadi, I.; Carter, A.S.; Subramanian, V.; Dillmann, I.; et al. Computational approaches identify a transcriptomic fingerprint of drug-induced structural cardiotoxicity. *Cell Biol. Toxicol.* **2024**, *40*, 50. [[CrossRef](#)]
90. Zhu, Z.; Wu, R.; Luo, M.; Zeng, L.; Zhang, D.; Hu, N.; Hu, Y.; Li, Y. Two-Dimensional Deep Learning Frameworks for Drug-Induced Cardiotoxicity Detection. *ACS Sens.* **2024**, *9*, 3316–3326. [[CrossRef](#)]
91. Huang, C.; Pei, J.; Li, D.; Liu, T.; Li, Z.; Zhang, G.; Chen, R.; Xu, X.; Li, B.; Lian, Z.; et al. Analysis and Validation of Critical Signatures and Immune Cell Infiltration Characteristics in Doxorubicin-Induced Cardiotoxicity by Integrating Bioinformatics and Machine Learning. *J. Inflamm. Res.* **2024**, *17*, 669–685. [[CrossRef](#)] [[PubMed](#)]
92. Long, M.; Alnoury, M.; Udupa, J.K.; Tong, Y.; Wu, C.; Poole, N.; Mannikeri, S.; Ky, B.; Feigenberg, S.J.; Zou, J.W.; et al. Prediction of Radiation Therapy Induced Cardiovascular Toxicity from Pretreatment CT Images in Patients with Thoracic Malignancy via an Optimal Biomarker Approach. *Acad. Radiol.* **2025**, *32*, 1895–1905. [[CrossRef](#)]
93. Zhao, X.; Sun, Y.; Zhang, R.; Chen, Z.; Hua, Y.; Zhang, P.; Guo, H.; Cui, X.; Huang, X.; Li, X. Machine Learning Modeling and Insights into the Structural Characteristics of Drug-Induced Neurotoxicity. *J. Chem. Inf. Model.* **2022**, *62*, 6035–6045. [[CrossRef](#)]
94. Pang, X.; He, X.; Yang, Y.; Wang, L.; Sun, Y.; Cao, H.; Liang, Y. NeuTox 2.0: A hybrid deep learning architecture for screening potential neurotoxicity of chemicals based on multimodal feature fusion. *Environ. Int.* **2025**, *195*, 109244. [[CrossRef](#)]
95. He, X.; Yang, Z.; Wang, L.; Sun, Y.; Cao, H.; Liang, Y. NeuTox: A weighted ensemble model for screening potential neuronal cytotoxicity of chemicals based on various types of molecular representations. *J. Hazard. Mater.* **2024**, *465*, 133443. [[CrossRef](#)]

96. Bercu, J.; Trejo-Martin, A.; Chen, C.; Schuler, M.; Cheung, J.; Cheairs, T.; Lynch, A.M.; Thomas, D.; Czich, A.; Atrakchi, A.; et al. HESI GTTC ring trial: Concordance between Ames and rodent carcinogenicity outcomes for N-nitrosamines (NAs) with rat and hamster metabolic conditions. *Regul. Toxicol. Pharmacol.* **2025**, *161*, 105835. [[CrossRef](#)]
97. Rane, R.; Satpute, B.; Kumar, D.; Suryawanshi, M.; Prabhune, A.G.; Gawade, B.; Mahajan, A.; Pawar, A.; Sakat, S. Mutagenic and genotoxic in silico QSAR prediction of dimer impurity of gliflozins; canagliflozin, dapagliflozin, and empagliflozin and in vitro evaluation by Ames and micronucleus test. *Drug Chem. Toxicol.* **2025**, *48*, 416–425. [[CrossRef](#)] [[PubMed](#)]
98. Li, Z.; Ma, W.; Ouyang, Y.; Cheng, P.; Zhong, D. Prediction of trihalomethane occurrence and cancer risk using interpretable machine learning and virtual data augmentation. *J. Hazard. Mater.* **2025**, *494*, 138697. [[CrossRef](#)] [[PubMed](#)]
99. Huang, Z.; Lou, S.; Wang, H.; Li, W.; Liu, G.; Tang, Y. AttentiveSkin: To Predict Skin Corrosion/Irritation Potentials of Chemicals via Explainable Machine Learning Methods. *Chem. Res. Toxicol.* **2024**, *37*, 361–373. [[CrossRef](#)]
100. Di, P.; Zheng, M.; Yang, T.; Chen, G.; Ren, J.; Li, X.; Jiang, H. Prediction of serious eye damage or eye irritation potential of compounds via consensus labelling models and active learning models based on uncertainty strategies. *Food Chem. Toxicol.* **2022**, *169*, 113420. [[CrossRef](#)]
101. Lynch, C.; Sakamuru, S.; Ooka, M.; Huang, R.; Klumpp-Thomas, C.; Shinn, P.; Gerhold, D.; Rossoshek, A.; Michael, S.; Casey, W.; et al. High-Throughput Screening to Advance In Vitro Toxicology: Accomplishments, Challenges, and Future Directions. *Annu. Rev. Pharmacol. Toxicol.* **2024**, *64*, 191–209. [[CrossRef](#)] [[PubMed](#)]
102. Dou, B.; Zhu, Z.; Merkurjev, E.; Ke, L.; Chen, L.; Jiang, J.; Zhu, Y.; Liu, J.; Zhang, B.; Wei, G.W. Machine Learning Methods for Small Data Challenges in Molecular Science. *Chem. Rev.* **2023**, *123*, 8736–8780. [[CrossRef](#)] [[PubMed](#)]
103. Luo, Y.S.; Ying, R.Y.; Chen, X.T.; Yeh, Y.J.; Wei, C.C.; Chan, C.C. Integrating high-throughput phenotypic profiling and transcriptomic analyses to predict the hepatosteatosis effects induced by per- and polyfluoroalkyl substances. *J. Hazard. Mater.* **2024**, *469*, 133891. [[CrossRef](#)]
104. Laban, B.B.; Ralevic, U.; Leskovic, A.; Petrovic, S.; Stoiljkovic, M.; Rosic, M.; Nikezic, A.V. Spectral and cytotoxicity studies of hybrid silver nanoparticles in human lymphocytes. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2025**, *340*, 126360. [[CrossRef](#)]
105. Mukhopadhyay, D.; Cocco, P.; Orru, S.; Cherchi, R.; De Matteis, S. The role of MicroRNAs as early biomarkers of asbestos-related lung cancer: A systematic review and meta-analysis. *Pulmonology* **2025**, *31*, 2416792. [[CrossRef](#)] [[PubMed](#)]
106. Zhou, X.; Liang, Z.; Zhang, G. Using explainable machine learning to investigate the relationship between childhood maltreatment, positive psychological traits, and CPTSD symptoms. *Eur. J. Psychotraumatol.* **2025**, *16*, 2455800. [[CrossRef](#)]
107. Son, A.; Park, J.; Kim, W.; Yoon, Y.; Lee, S.; Ji, J.; Kim, H. Recent Advances in Omics, Computational Models, and Advanced Screening Methods for Drug Safety and Efficacy. *Toxics* **2024**, *12*, 822. [[CrossRef](#)]
108. Nelson, C.P.; Brown, P.; Fitzpatrick, S.; Ford, K.A.; Howard, P.C.; MacGill, T.; Margerrison, E.E.C.; O'Shaughnessy, J.; Patterson, T.A.; Raghuwanshi, R.; et al. Advancing alternative methods to reduce animal testing. *Science* **2024**, *386*, 724–726. [[CrossRef](#)]
109. Bai, C.; Wu, L.; Li, R.; Cao, Y.; He, S.; Bo, X. Machine Learning-Enabled Drug-Induced Toxicity Prediction. *Adv. Sci.* **2025**, *12*, e2413405. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.