Environmental Science & Technology

pubs.acs.org/est

Perspective

In Silico Frontiers Shaping the Next Generation of Transformation Product Prediction and Toxicological Assessment

Published as part of Environmental Science & Technology special issue "Nobel Symposium 2025: The Future of Chemical Safety and Sustainable Materials Chemistry".

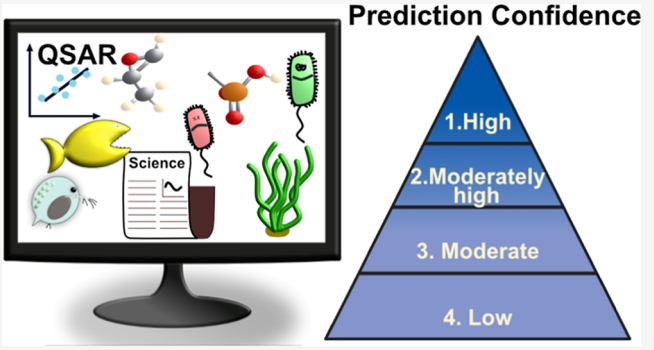
Paul Löffler,* Emma L. Schymanski, Henning Henschel, and Foon Yin Lai*

Cite This: https://doi.org/10.1021/acs.est.5c06790

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: The characterization of transformation products (TPs) is crucial for understanding chemical fate and potential environmental hazards. TPs form through (a)biotic processes and can be detected in environmental concentrations comparable to or even exceeding their parent compounds, indicating toxicological relevance. However, identifying them is challenging due to the complexity of transformation processes and insufficient data. *In silico* methods for predicting TP formation and toxicity are efficient and support prioritization for chemical risk assessment, yet require sufficient data for improved results. This perspective article explores the role of computational approaches in assessing TPs and their potential effects, including rule-based models, machine learning-based methods, and QSAR-based toxicity predictions, focusing on openly available tools. While integrating these approaches into computational workflows can support regulatory decision-making and prioritization strategies, predictive models can face limitations related to applicability domains, data biases, and mechanistic uncertainties. To better communicate the results of *in silico* predictions, a framework of four distinct levels of confidence is proposed to support the integration of TP prediction and toxicity assessment into computational pipelines. This article highlights current advances, challenges, and future directions in applying *in silico* methodologies for TP evaluation, emphasizing the need for more data and expert interpretation to enhance model reliability and regulatory applicability.



KEYWORDS: *environmental fate, computational (eco)toxicology, chemical prioritization, risk assessment, QSAR modeling, rule-based models, machine learning, organic micropollutants*

1. INTRODUCTION

The importance of characterizing transformation products (TPs) potentially affecting the receiving aquatic environments has been increasingly emphasized,[1−4] with many TPs found in similar or even higher environmental concentrations than their respective parent compound.[5−7] For example, Kołecka et al. quantified two diclofenac TPs in effluent wastewater with concentration levels almost double than diclofenac itself.[8] However, discovering all possible TPs is challenging. Several Organization for Economic Co-operation and Development (OECD) guidelines exist to investigate environmental (e.g., photo, microbial) transformation of chemicals in aquatic ecosystems.[9−12] This perspective considers TPs from multiple transformation pathways, including abiotic processes such as photolysis or water treatment, and biotic processes such as environmental biotransformation and human metabolism. TPs formed within living organisms (i.e., metabolites or biotransformation products) can be identified via *in vitro* or *in vivo* methods. The former involves exposure of a chemical to specific enzymes in laboratory-scale experiments, while the later refers to the analysis of biological matrices, such as blood, tissue, or excreta, following exposure to a chemical. Complicating factors in these methods include ethical considerations and the variability across different organisms and environmental contexts.[13−19] TPs formed through abiotic reactions such as photolysis or treatment processes can be determined through laboratory experiments or pilot plants, with sophisticated setups.[20−26] The analytical method of choice for identifying and discovering new TPs is high-

Received: May 20, 2025
Revised: August 13, 2025
Accepted: August 14, 2025

resolution mass spectrometry (HRMS), generating extensive datasets that require careful investigation to accurately identify each TP, with many features remaining unidentified or only tentatively identified.[3,27,28] Several TPs have been shown to contribute to the overall hazard and risk profile in the environment.[29−36] For example, fluoxetine,[37] propranolol,[38] and acyclovir[3] TPs have been suggested to exhibit (eco-)toxicological effects. Recently, 6PPD-quinone, the TP of the tire additive 6PPD (4-*N*-(4-methylpentan-2-yl)-1-*N*-phenylbenzene-1,4-diamine) that can enter environmental waters through for example urban runoff, was shown to exhibit toxic effects to multiple fish species, with toxicity levels several orders of magnitude higher than 6PPD itself.[30,39−43] Given the established contribution of several TPs to the overall hazard and risk profile of environmental samples, a holistic risk assessment aims at covering as much of the chemical space as possible. However, it is neither practical nor realistic to assess risks of all potential chemicals and their TPs individually through HRMS and ecotoxicological studies.

Combining chemical with effect-based methods and *in silico* approaches has been suggested to investigate combined effects and mechanisms of toxicity.[44−46] *In silico* methodologies can help to fill knowledge gaps and support screening or prioritization. Computational approaches can predict how chemicals would behave in the environment and their potential toxic effects, including quantitative structure activity relationships (QSARs) and read-across methods.[47,48] Additionally, molecular docking and molecular dynamics simulations, widely used in medicinal chemistry, are increasingly considered for chemical safety assessments, offering potential insights into toxic action mechanisms.[46] Comprehensive workflows can now predict TPs and key toxicological endpoints from just the initial chemical structure. Such approaches could serve as essential safety measures, for example, in early assessment stages for regulatory and drug design purposes, enabling more informed decision-making in chemical production. Additionally, these methodologies allow for the integration of TP assessments, aiding environmental scientists and other stakeholders in managing chemical impacts effectively.

This perspective article explores how *in silico* methodologies can enhance the risk assessment process for TPs in order to facilitate the development of computational workflows that integrate TP formation and toxicity assessments. This could be beneficial to various fields, including pharmaceutical development and environmental sciences, by enabling proactive evaluations of chemical safety and environmental impacts. The motivation stems from recent recommendations within the scientific community for early integration of persistence and toxicity measures into management frameworks to implement a more proactive approach.[49−51] This article focuses on broadly applicable open access *in silico* approaches for predicting TPs and toxicological impacts. Tools are compared based on their functionality, input requirements, applicability domain, interpretability, and validation strategies. This work also highlights emerging computational approaches, current challenges, and research needs in TP prediction and toxicological assessment.

2. FOUNDATIONS OF PREDICTIVE APPROACHES

There are two primary computational approaches: rule-based models and machine learning-based models, each with strengths and limitations, offering complementary insights into chemical behavior and risks.

2.1. Rule-Based Models. Rule-based models are grounded in mechanistic evidence derived from experimental studies. They rely on predefined rules or structural alerts, molecular substructures or patterns associated with specific biological activities, transformations, or toxicological endpoints. In TP prediction, rule-based models apply expert-curated reaction rules to forecast transformations such as hydroxylation or oxidation. In toxicology, the presence of a structural alert, such as a nitro group linked to mutagenicity,[52] can serve as indicator for hazard identification. The interpretability of rule-based models is one of their key strengths, as they are built on well-defined reaction pathways or mechanistic insights. However, they are inherently constrained by the width and depth of their underlying libraries. This means they can only predict behaviors and transformations/mode of actions that have already been characterized, limiting their utility for novel chemicals or uncharted mechanisms.

2.2. Machine Learning Models. Machine learning (ML) models are data-driven and particularly effective in capturing complex, nonlinear relationships. By analyzing large datasets of chemical properties, structures, and biological activities, these models can uncover patterns and make predictions that extend beyond existing mechanistic knowledge.[53] In TP prediction, ML algorithms can predict potential transformation pathways based on chemical descriptors and environmental factors. In toxicological assessment, ML models can estimate effects like bioaccumulation or endocrine activity by learning from extensive experimental datasets. While ML models are powerful and flexible, their reliability depends on the quality, diversity, and size of the training datasets. They also face challenges like overfitting, where the model performs well on training data but poorly on unseen data. Additionally, the black-box nature of many ML methods can hinder interpretability, making it difficult to trace predictions back to mechanistic insights.

2.3. Integration and Complementarity. Rule-based and ML models are not mutually exclusive but complementary. Workflows and approaches that integrate both these approaches combine the reliability of expert knowledge with the adaptability of data-driven insights. QSAR models serve as a bridge between rule-based and ML approaches, as they can be developed using expert-defined descriptors rooted in mechanistic knowledge or trained on large datasets using statistical learning methods. Similarly, read-across approaches, which involve predicting properties of a target chemical using data from structurally similar, well-studied analogues, are increasingly enhanced by ML to improve predictive accuracy.[54,55] This combined approach forms the foundation of predictive methodologies discussed in the following sections, illustrating how these techniques are applied.

3. FINDING DATA ON KNOWN TRANSFORMATION PRODUCTS

Datasets of known TPs are the starting point for most investigations and form the basis for developing rule-based and ML approaches discussed above. Systematic literature searching (e.g., predefining specific search strings and using multiple scientific databases) usually results in a large number of articles that need to be screened. Multiple text-mining tools[56−59] assist and facilitate this work, including chemical data extraction pipelines.[60−62] ShinyTPs was specifically designed to curate TP information derived from text-mining of hand-selected text snippets integrated within PubChem.[62]

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

With increased contribution to and awareness of open access TP resources, such as enviPath[63,64] and suspect lists on the NORMAN Suspect List Exchange (NORMAN-SLE),[65] screening existing databases[66] or shared suspect lists for TPs[67−82] has become more common. Several lists with parent-TP mappings on the NORMAN-SLE[65] have been mapped up into transformations templates,[83] added into PubChem in the "Transformations" section and archived as an (updatable) data set on Zenodo.[66] This enables both public display (in PubChem) to raise awareness of the data, and integration into TP identification workflows, such as those integrated within patRoon.[21,84,85] This collaborative community effort currently includes 9152 unique reactions involving 9267 unique compounds. Of the chemicals included, 3724 are classified as parents and 7331 as TPs (some are both parent and TPs in different reactions). Although these numbers have grown considerably in the last years and are now triple what was used to train BioTransformer[86,87] (detailed further below), this is still a tiny fraction (<0.1%) of the currently >131 000 compounds in the NORMAN-SLE,[65] and an even smaller fraction (<0.0001%) of the chemicals in PubChem. The lack of sufficiently documented open data on TPs is a huge challenge for establishing reliable computational methods, as the current knowledge focuses on only certain chemical classes in great detail, yet does not cover many other classes that are known to be present in these databases.

While it is feasible that large language models (LLMs), such as ChatGPT, can be prompted to propose lists of possible TPs, they should be treated with caution, as their outputs are not based on curated chemical reaction rules or mechanistic understanding, and assessing their applicability domain is currently not feasible. To date, systematic exploration or scientific validation of LLMs for TP prediction is lacking. In-depth analysis and prediction using LLMs is therefore not recommended, as they can often generate plausible-sounding but false or unverifiable information.[88,89] In contrast, databases documenting known TP reactions offer a higher level of reliability and transparency, as they provide carefully curated data by experts following strict criteria for data inclusion and referencing protocols for verification, ensuring a more trustworthy source of information.

4. PREDICTION OF TRANSFORMATION PRODUCTS

In silico strategies that predict TPs using expert knowledge or pattern recognition for the creation of suspect lists for improved screening in HRMS experiments have gained attention.[28] These computational tools are valued for their ability to generate novel chemical structures, whether plausible or not. The *in silico* TP prediction tools discussed in this work incorporate a comprehensive array of underlying transformation rules and models, tailored for diverse processes such as phase I or phase II metabolism, and environmental microbial degradation. With increasing attention to advanced treatment technologies, it is feasible that these approaches could be expanded to cover such transformation reactions as more data on TPs from advanced treatment processes becomes available. To support these advancements, it is crucial that researchers share experimental data on transformation reactions, to enhance model development and validation. The ACS author guidelines for several environmental journals have recently been updated to provide some instructions and suggestions to authors how to share this information.[90] Unless otherwise specified, the tools discussed below are limited to organic compounds under ∼1000−1500 Da, and do not support polymers, nanomaterials, or highly fluorinated substances due to a lack of representative training data or rules.

BioTransformer, an open source tool, includes eight models of metabolic transformation prediction, including phase I (cytochrome P450), promiscuous enzymatic, phase II, human gut microbial, environmental microbial transformations and different combinations of the above known as *AllHuman*, *SuperBio* and *MultiBio*.[86,87] Users can submit molecular structures as Simplified Molecular-Input Line-Entry System (SMILES), a line notation describing chemical structures, or as a Structured Data File (SDF), a standard format for storing molecule structure information and associated data. BioTransformer is available as command-line tool and through a web server at www.biotransformer.ca. While it supports batch processing of chemicals, it does not allow for batch mode across multiple models. However, this limitation can be overcome using the command line version and a bash script (example file and explanation can be found here: https://github.com/paloeffler/biotrans_multiprompt) that loops over all the models of interest. The web tool outputs an interactive table of the predicted TPs. An example of antimicrobial TPs generated via BioTransformer and the mentioned script is published online in NORMAN suspect list S114.[82] BioTransformer integrates rule-based and ML approaches, and its underlying data, including biotransformation rules and a curated database (MetXBioDB), are openly accessible through a web service, as a downloadable Java Library[91] and on the NORMAN-SLE.[74] A major update, BioTransformer 4.0, is expected soon but is not officially released at the time of writing. It introduces over 130 new reaction rules, a validation module that filters unrealistic metabolites based on similarity to known human metabolites, and a new abiotic metabolism module covering photolysis, chlorination, and ozonation reactions, partly derived from the CTS database. In the environmental metabolism module, the update improves SMIRKS string handling and fixes incorrect transformation rules that previously produced invalid metabolites.

A second option offering a variety of transformation algorithms is the Reaction Pathway Simulator module in the Chemical Transformation Simulator (CTS) by the U.S. EPA.[92] It integrates various tools, such as EPISuite, the Toxicity Estimation Software Tool (T.E.S.T.), ChemAxon and OPEn structure−activity/property Relationship App (OPERA). CTS offers flexible input options (Name, SMILES, CAS, sketcher input). CTS employs defined reaction libraries that include generalized reaction schemes, specifying how a molecular fragment is modified by a particular transformation process. When a molecule is submitted, CTS compares its structure to the reactant side of these schemes in the libraries. If a match is found, the tool modifies the matched fragment while leaving the rest of the molecule unchanged. This mechanism is not unique to CTS, but rather the general principle of rule-based approaches. CTS prioritizes predicted TPs by ranking them based on transformation rates reported in scientific literature. Currently, CTS provides reaction libraries for abiotic hydrolysis, abiotic reduction, direct photolysis, spontaneous reactions (e.g., dehydration of geminal diols), human phase I metabolism, and both environmental and metabolic reactions of per- and polyfluoroalkyl substances (PFAS). Each reaction library includes schematic reactions and references to the scientific rules underlying the predictions. Additionally, CTS offers integration with other tools such as BioTransformer and

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Table 1. Overview of the *In Silico* Tools Described in This Article, Their Included Models/Endpoints, Data Accessibility and Applicability Domain Estimation (Further Details Are Given in the Main Text)

Tool	Main focus	Included models	Training dataset accessible	Applicability domain provided
EPISuite[97]	physicochemical properties, ecotoxicology	multiple QSARs and ECOSAR	limited	not for all models
ToxTree[100]	toxicological hazard screening	cramer rules, verhaar scheme, Benigni/Bossa rules	yes	rule-based
T.E.S.T.[101]	ecotoxicology, human toxicity	QSARs	yes (ECOTOX database)	yes
OPERA[102]	physicochemical properties, human endocrine activity	CERAPP, CoMPARA, CATMoS	yes	yes
VEGA-QSAR[103]	physicochemical properties, ecotoxicology, toxicology, environmental fate	>100 models from CAESAR, OPERA, ECOSAR, etc.	yes	yes
TRIDENT[104]	ecotoxicology	deep learning transformer model	yes (Github)	yes
NR-ToxPred[105]	human endocrine activity	9 receptor models	yes	yes

enviPath Pathway Predictions, accessible through their respective APIs. While CTS has a GitHub repository (https://github.com/quanted/cts_app), much of its code relies on licensed software, limiting the creation of a fully independent clone. However, users can incorporate CTS into individual workflows via its REST API (https://qed.epa.gov/cts/rest/).

Another option to present here for TP prediction is the EAWAG-Biocatalysis/Biodegradation Database (BBD) Pathway Prediction System (PPS), which is also a rule-based, substructure searching, and atom-to-atom mapping prediction algorithm based on the biodegradation/biocatalysis database of the University of Minnesota.[93,94] The 249 biotransformation rules are publicly accessible (http://eawag-bbd.ethz.ch/servlets/pageservlet?ptype=allrules) and typically include a scientific reference for each reaction. Reaction rules are also prioritized based on likelihood assigned by an expert panel to each reaction. This ranges from very likely and likely (e.g., spontaneous hydrolysis in water), possible for reactions that are common but not certain to occur in every system (e.g., transformation of a secondary alcohol to a ketone), to unlikely and very unlikely for reactions only very rarely catalyzed in bacteria or fungi (e.g., reductive dehalogenation). The BBD-PPS terminate its prediction once certain small compounds are reached (http://eawag-bbd.ethz.ch/servlets/pageservlet?ptype=termcompsview). These terminal compounds include two categories: (1) small, readily degraded molecules that do not undergo further transformation, and (2) dead-end compounds, often larger or halogenated, that are known to persist in the environment due to their resistance to microbial degradation. If a compound in category (1) is encountered, its biodegradation is not predicted further, but instead a link to a relevant Kyoto Encyclopedia of Genes and Genomes (KEGG)[95] metabolic pathway is given. For compounds in category (2), no further transformation or KEGG pathway is offered. enviPath (envipath.org) expands the capabilities of the BBD-PPS with updated and more comprehensive reaction rules, an enhanced user interface, and integrated links to additional biochemical pathway databases, offering a more robust and user-friendly experience for exploring biotransformation pathways.[63] While BBD-PPS advised caution with molecules over 1000 Da and excluded PFAS and highly fluorinated chemicals due to limited rule coverage, enviPath addresses these limitations. A recent addition is a dedicated PFAS (per- and polyfluoroalkyl substances) package,[96] which includes curated microbial transformation pathways and trained reaction rules for selected fluorinated precursors. This targeted effort extends enviPath's predictive reach toward highly persistent and environmentally relevant contaminants. Furthermore, enviPath's open access database supports user contributions, enabling the continuous evolution of its predictive capabilities and the inclusion of diverse environmental conditions. This approach broadens the scope of chemicals that can be analyzed and improves the selectivity and reliability of the predictions.

Recently, the open-source platform patRoon,[21,84,85] integrated several of these predictive techniques into a pipeline connecting *in silico* predictions with HRMS data. Alongside the tools already discussed, patRoon includes the PubChem/NORMAN-SLE transformation datasets as well,[65] allowing users to systematically screen and annotate known and predicted TPs in their experimental data. This modular and extensible workflow enables researchers to efficiently prioritize and confirm TPs. Functionality for photolysis-related TP prediction and screening was added in 2025, further expanding patRoon's ability to capture both biotic and abiotic transformation pathways.[21] Through this integration, patRoon enhances the efficiency, reproducibility, and transparency of nontarget and suspect screening workflows.

As described above, enviPath is a highly curated predictive system specifically for environmental use cases, whereas CTS and BioTransformer offer environmental and additional metabolism functions. CTS also integrates abiotic reactions covering advanced treatment processes (functionality that is currently being developed in BioTransformer). Both CTS and BioTransformer integrate enviPath, while patRoon (a HRMS processing software) integrates all approaches and more. Thus, each approach offers significant overlap and the choice of which is the best in various scenarios may come down to user preferences.

5. TOXICOLOGICAL ASSESSMENT TOOLS

Unless otherwise specified, all tools discussed in this section (Table 1) are designed for organic compounds with well-defined molecular structures and do not support mixtures, substances of unknown or variable composition, nanomaterials, or polymers. These are general limitations of current QSAR and ML models due to the lack of consistent structural representation and training data for such complex substances.

A widely recognized predictive toxicity tool is the Estimation Program Interface, or EPISuite.[97] EPISuite integrates various models to estimate physicochemical properties and the Ecological Structure Activity Relationships (ECOSAR) predictive models, which are also available separately. ECOSAR models estimate aquatic ecotoxicity based on equations derived from experimental data, allowing for the evaluation

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

of several endpoints across multiple organisms within the aquatic food chain. These include green algae (72 or 96 h tests), *Daphnia* (48 h tests), and fish (96 h tests) for both acute lethality and chronic values. The user interface supports batch mode processing. While EPISuite results are validated internally, limited availability of the training and validation datasets hamper independent assessment of the applicability domains (Table 1). Recent studies highlighted limitations for phytotoxins[98] and those with atypical functional groups, particularly for fluorinated and phosphorus-containing compounds.[99]

A free open-source rule-based tool to predict the toxicological hazard of chemicals is ToxTree.[100] It applies various decision tree models incorporated into the concept of threshold of toxicological concern to assess the so-called Cramer class of a chemical substance to estimate its relative toxic hazard. ToxTree evaluates chemical structures against a set of predefined rules or structural alerts to determine potential hazards, which is useful for initial hazard assessment in chemical safety evaluation. ToxTree offers multiple classification schemes, including Cramer decision tree for oral toxicity classification, Verhaar scheme for mode of toxic action of organic chemicals, Benigni/Bossa rule-based mutagenicity and carcinogenicity alerts. The tool provides transparent and interpretable results, as each classification follows explicit mechanistically relevant rules. ToxTree supports batch processing and accepts SMILES, MOL, and SDF files as input formats.

The Toxicity Estimation Software Tool (T.E.S.T.) incorporates the Computer Assisted Evaluation of industrial chemical Substances According to Regulations (CAESAR) model for developmental toxicity as well as carcinogenicity and mutagenicity models, also implemented in VEGA-QSAR. The open-access tool also incorporates models for the prediction of endpoints for fathead minnow LC_{50} (96 h), *Daphnia magna* LC_{50} (48 h), *tetrahymena pyroformis* IGC_{50}, oral rat toxicity (LD_{50}), and bioaccumulation factor for fish.[106−110] T.E.S.T. uses several ML models along with conventional QSAR methods and accepts CAS, SMILES, name, InChI, InChIKey, DTXSID, or sketcher input. Batch mode processing is supported (txt, SMILES, SDF). Compounds must have defined structures and fall within the model's molecular weight range (≤2000 Da). The outputs are offered in different formats (csv, excel or html). The batch mode processes multiple chemicals for only a single end point at one time. Model specific validation results for T.E.S.T. are documented in the User's Guide, while all experimental toxicity data used for model development originates from the publicly available ECOTOX database, allowing for independent evaluation and further analysis.

The OPEn qsaR App (OPERA) includes predictions for estrogenic activity from the Collaborative Estrogen Receptor Activity Prediction Project (CERAPP),[111] Androgenic activity from the Collaborative Modeling project for Androgen Receptor Activity (CoMPARA),[112] as well as the acute oral systematic toxicity from the Collaborative Acute Toxicity Modeling Suite (CATMoS),[113] and predictions of physicochemical properties such as acid dissociation constant, octanol−water partitioning coefficient and distribution constant for nonionizable compounds.[114−116] OPERA is open source (https://github.com/kmansouri/OPERA) and can be used locally with or without graphical user interface. It is included in several open resources, including the U.S. EPA CompTox Chemicals Dashboard[117] and as extension in the QSAR Toolbox.[118,119] OPERA allows batch mode processing with various input formats (SMILES, SDF, MOL, CASRN, DTXSID, DTXCID, InChIKey) and returns a list of molecule IDs, predictions, the applicability domain and an accuracy assessment.[102,120] One of OPERA's key strengths is its applicability domain assessment, based on structural similarity measures, leverage-based methods, and distance-to-model calculations, to assess how closely a given compound aligns with its training data set.

VEGA-QSAR is an open-access tool integrating over 100 predictive models, combining various QSAR-based toxicological, environmental, and physicochemical assessments. It incorporates models from CAESAR,[121,122] OPERA, EPI Suite,[102,123,124] and others,[100,125] supporting regulatory and environmental applications. VEGA has put emphasis on ensuring that the models generate transparent and reproducible results, providing model guides, test and training datasets accessible in the standalone application (Figure 1), facilitating

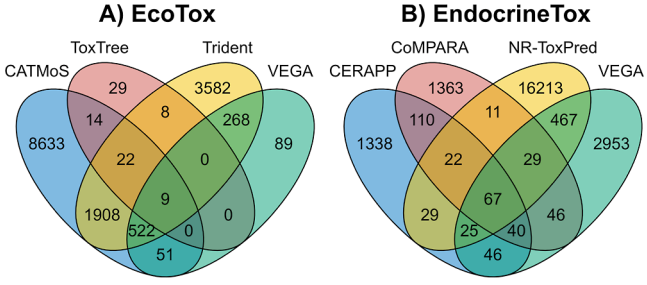


Figure 1. Number of compounds included in the training and test datasets for (A) ecotoxicological endpoints (EcoTox) and (B) endocrine endpoints (EndocrineTox). For VEGA in A), the datasets used were fish acute LC_{50} SarPy/IRFMN, *Daphnia magna* LC_{50} IRFMN, and algae acute EC_{50} IRFMN. For VEGA in B), the datasets used were androgen receptor-mediated effect (IRFMN/CoMPARA), estrogen receptor-mediated effect (IRFMN/CERAPP), and estrogen receptor relative binding affinity (IRFMN). Datasets were merged using SMILES and CAS numbers when available.

screening of these datasets and checking the applicability of the respective model. It supports different standard formats used in the chemical domain, including SMILES and SDF. Batch mode is available, including multiple model selection. VEGA can also be used for read-across approaches without involving QSAR models.[126]

A recent model for ecotoxicological end point prediction is the deep learning model TRIDENT,[104] which is based on the transformer architecture. TRIDENT predicts two toxicity endpoints, EC_{50} and EC_{10}, for three species groups (algae, aquatic invertebrates and fish) and a variety of effects. The web-service version uses SMILES (https://trident.serve.scilifelab.se/) and allows, depending on the combination of end point and species group, predictions for mortality, intoxication, population, reproduction, and growth. The code, full model and data set used to develop the model, consisting of almost 150 000 experimental data for 6657 unique chemicals (Figure 1), are available online (https://github.com/StyrbjornKall/TRIDENT). The training data set includes a large fraction of charged chemicals (~25%), including inorganic compounds such as NiF_2, $FeCl_3$, Fe_2O_3, $PbSO_4$ and PdO. While most tools exclude such compounds, TRIDENT's training data include a number of organometallics

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

like hydroxy-methylmercury, expanding its coverage slightly beyond typical mode. TRIDENT outperformed three existing models (ECOSAR, VEGA, and T.E.S.T.) for most endpoints, except algae EC_{50}.[104]

In addition to OPERA, the ML model NR-ToxPred offers *in silico* predictions of endocrine activity by assessing ligand binding to nine human nuclear receptors (e.g., androgen, estrogen α/β, progesterone). Based on a public data set of ~15,000 entries (Figure 1), the model provides binary predictions (active/inactive, binding/nonbinding) along with sensitivity, specificity, and applicability domain estimates using the Tanimoto similarity measure.[127] Unlike OPERA, NR-ToxPred does not distinguish between agonists and antagonists, lacks uncertainty quantification, and is limited to organic compounds. Although the model code is not public, the tool is accessible via a user-friendly web interface (http://nr-toxpred.cchem.berkeley.edu/) and supports batch prediction with CSV input and receptor binding site visualization.

There are numerous other toxicity prediction models available, targeting specific organisms, endpoints, or effects, as detailed elsewhere.[119,128−135] The online chemical modeling environment (OCHEM) can be used to run available models to screen compounds for structural alerts for (eco)toxicological endpoints, and also provides the opportunity to create new QSAR models based on the experimental data in the database.[136−139] Two research groups have recently developed algorithms to estimate ecotoxicity endpoints from HRMS fragment data.[140,141] Such approaches could facilitate chemical risk assessment from chemical screening data and provide further insights into mixture toxicity assessment. Additionally, conventional dose−response models may fall short in accounting for continuous low-level exposure or the specific toxicokinetic behavior of highly persistent or bioaccumulative substances.[142] For example, differences in compound distribution, such as accumulation in fatty tissues versus protein binding, can significantly affect internal exposure and toxicodynamics. The integration of pharmacokinetic-pharmacodynamic modeling, which assesses the relationship between chemical exposure and biological response over time, could enhance prediction accuracy by incorporating absorption, distribution, metabolism, and excretion dynamics. These models are particularly relevant for widespread contaminants and extremely persistent chemicals, where chronic exposure scenarios may be more representative of real-world environmental conditions. In cases where a hypothesis of the specific mode of toxic actions exists, this can be confirmed and its understanding deepened via *in silico* tools, such as molecular docking or molecular dynamic simulations with free energy perturbations, as discussed recently.[46] These techniques require more bioinformatics and command line skills than the previously described approaches, but could initiate the development of adverse outcome pathways and by that contribute for example to a computational ecotoxicity assay.[45]

6. REMARKS FOR FUTURE

In silico approaches for TP and toxicity predictions are beneficial to researchers and legislators in providing additional acquisition of toxicity-related information on TPs. Advances in ML and computational power have made it easier to develop predictive models; however, meaningful improvements in prediction accuracy depend on robust validation methods and well-defined criteria. While models are becoming more sophisticated, many suffer from overfitting, heavy bias, or poor generalizability due to for example limited and biased training datasets. A clear understanding of estimation methods and their appropriate application is therefore critical. Beyond ensuring alignment with best-practice guidelines,[143−146] we propose four distinct levels of confidence (Figure 2) to be reported for enhancing both interpretability and reliability of TP predictions.

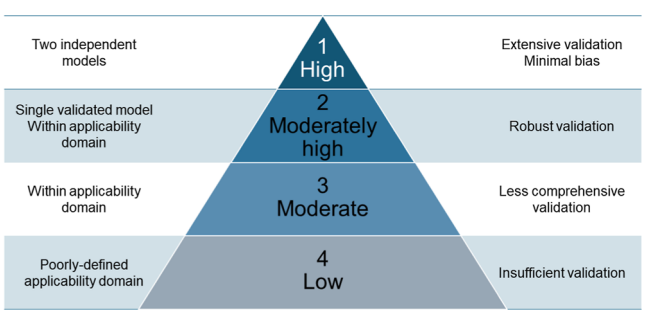


Figure 2. Schematic visualization of the confidence levels including defining criteria.

1. High confidence (validated and reliable)

Two or more independent models with well-defined applicability domains and extensive validation across diverse datasets. Minimal bias, strong generalization across chemical classes, and mechanistic support from rule-based models with literature backing up.

Example: Prediction of acute fish toxicity for 4-nitrophenol using VEGA-QSAR and TRIDENT. The compound falls within the applicability domain of both models and is included in their training datasets. This direct inclusion greatly enhances the reliability and confidence in the predicted toxicity values.

2. Moderately high confidence (reliable but less broadly validated)

Single validated model with a well-defined applicability domain, robust validation, and transparent methodology (e.g., public datasets). Rule-based models supported by mechanistic plausibility but lacking experimental confirmation for similar chemical compounds.

Example: Prediction of estrogen binding potential of bisphenol S using the OPERA platform (CEARPP model for estrogenicity). The prediction is within the model's applicability domain and supported by robust validation and clear mechanistic relevance. Although no experimental data for bisphenol S are present in the model's training data set, its close analogue bisphenol A is well represented, providing additional support and resulting in moderately high confidence in the prediction.

3. Moderate confidence (limited generalization)

Predictions within the applicability domain but with less comprehensive validation or uncertain generalization beyond specific datasets. Rule-based models relying on mechanistic assumptions but lacking empirical validation for the relevant chemical class.

Example: Prediction of acute Daphnia toxicity for ciprofloxacin using the VEGA-QSAR model is of moderate confidence. While the compound's broad structure may be technically within the model's applicability domain, ciprofloxacin and related fluoroquinolone antibiotics are not represented in the VEGA training set, and the model has not been comprehensively validated for this chemical

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

class. Therefore, there is uncertainty in the prediction's reliability for antibiotics with ionizable and zwitterionic properties.

4. Low confidence (uncertain or limited reliability)

Predictions from models with poorly defined applicability domains, insufficient validation, or high uncertainty in extrapolation.

Example: Prediction of acute algal toxicity for novel silicon-containing compound using T.E.S.T model. However, because organosilicons are not represented in the training data and the applicability domain for this class is poorly defined, the reliability of the prediction is considered low confidence.

Following the European Food Safety Authority (EFSA) guidelines, the use of two independent QSAR models confirming predictions is recommended,[147,148] where independence refers to differing training datasets or algorithms (rule-based *vs* statistical). Both models should be of high to moderate-high confidence. Most models do not account for mixture toxicity effects (e.g., additive or synergistic effects of chemicals).[149] Furthermore, environmental conditions can vary and should be considered for ionic and ionizable chemicals, as these factors can govern e.g., the partitioning in environmental systems.[150,151] The validation of most predictive toxicology models using novel compounds (not included in any test or training data set) with different modes of action is of high interest to experimentally validate accuracy and precision of the models.

While this article highlights the potential for computational TP and toxicity prediction methodologies to support research and enhance risk assessments of TPs, predictive reliability remains variable across different chemical classes due to uneven data coverage. A concerted community effort on generating and sharing relevant data for greater portions of the "chemical space", rather than generating yet more data for compounds very similar to existing data, would help expand the applicability domains—and thus increasing the usefulness of these computational approaches immensely. Additionally, TPs formed during water treatment processes (e.g., advanced oxidation processes like ozonation) are gaining attention, especially in light of the recast EU wastewater treatment directive (EU 2024/3019).[152] Despite their growing environmental relevance, these treatment-derived TPs are often underrepresented or unsupported in current *in silico* tools, although recent developments are striving to cover this gap. Expanding the underlying experimental data collections as well as model rules/coverage to include these TPs would help align computational assessments more closely with real-world transformation pathways and support regulatory needs.

AUTHOR INFORMATION

Corresponding Authors

Paul Löffler − *Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences (SLU), SE-750 07 Uppsala, Sweden;* orcid.org/0000-0003-1959-0752; Email: paul.loffler@slu.se

Foon Yin Lai − *Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences (SLU), SE-750 07 Uppsala, Sweden*; Email: foonyin.lai@slu.se

Authors

Emma L. Schymanski − *Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 4367 Belvaux, Luxembourg;* orcid.org/0000-0001-6868-8145

Henning Henschel − *Department of Medicinal Chemistry, Uppsala University (UU), Uppsala SE-751 21, Sweden;* orcid.org/0000-0001-7196-661X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.5c06790

Notes

The authors declare no competing financial interest.

Biographies

Paul Löffler is a PhD candidate at the Swedish University of Agricultural Sciences (SLU) investigating the impact of antimicrobial transformation products on aquatic environments. Holding a bachelor's degree in chemistry from the University of Stuttgart and a master's degree in ecotoxicology from the University Koblenz-Landau, he integrates his expertise together with multidisciplinary colleagues from for example medicinal and computational chemistry to enhance *in silico* methodologies.

Foon Yin Lai is a senior lecturer in the field of Analytical and Environmental Chemistry. Her group is researching on chemical use in society (wastewater-based epidemiology), water pollution, source elucidation and (waste)water reuse related to emerging contaminants. In these topics, her group develops new analytical methodology for chemical detection and also workflows with *in silico* tools and new approaches for prioritizing chemicals of concern and for chemical risk assessment. She is interested in studying transformation products and other chemicals associated with negative health effects, e.g., antimicrobial resistance and endocrine disruption. She obtained her Ph.D. in Environmental Forensic Chemistry from The University of Queensland (Australia) in 2014, and has been as an Associate Professor at the Swedish University of Agricultural Sciences (SLU, Sweden) since 2020.

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

■ ACKNOWLEDGMENTS

This work was funded by Swedish Research Council (project number: 2020-03675). The authors also thank Mikael Gustavsson for discussions on TRIDENT and Emma Palm (LCSB, University of Luxembourg) for various discussions related to this manuscript.

■ REFERENCES

(1) Richardson, S. D.; Ternes, T. A. Water Analysis: Emerging Contaminants and Current Issues. *Anal. Chem.* **2022**, *94* (1), 382−416.

(2) Yang, Y.; Zhang, X.; Jiang, J.; Han, J.; Li, W.; Li, X.; Yee Leung, K. M.; Snyder, S. A.; Alvarez, P. J. J. Which Micropollutants in Water Environments Deserve More Attention Globally? *Environ. Sci. Technol.* **2022**, *56* (1), 13−29.

(3) Löffler, P.; Escher, B. I.; Baduel, C.; Virta, M. P.; Lai, F. Y. Antimicrobial Transformation Products in the Aquatic Environment: Global Occurrence, Ecotoxicological Risks, and Potential of Antibiotic Resistance. *Environ. Sci. Technol.* **2023**, *57* (26), 9474−9494.

(4) Wang, X.; Yu, N.; Yang, J.; Jin, L.; Guo, H.; Shi, W.; Zhang, X.; Yang, L.; Yu, H.; Wei, S. Suspect and Non-Target Screening of Pesticides and Pharmaceuticals Transformation Products in Wastewater Using QTOF-MS. *Environ. Int.* **2020**, *137*, 105599.

(5) Boxall, A. B. A.; Sinclair, C. J.; Fenner, K.; Kolpin, D.; Maund, S. J. Peer Reviewed: When Synthetic Chemicals Degrade in the Environment. *Environ. Sci. Technol.* **2004**, *38* (19), 368A−375A.

(6) Transformation Products of Synthetic Chemicals in the Environment*The Handbook of Environmental Chemistry*; Boxall, A. B. A., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2009; Vol. *2P*.

(7) Escher, B. I.; Fenner, K. Recent Advances in Environmental Risk Assessment of Transformation Products. *Environ. Sci. Technol.* **2011**, *45* (9), 3835−3847.

(8) Kołecka, K.; Gajewska, M.; Stepnowski, P.; Caban, M. Spatial Distribution of Pharmaceuticals in Conventional Wastewater Treatment Plant with Sludge Treatment Reed Beds Technology. *Sci. Total Environ.* **2019**, *647*, 149−157.

(9) OECD *Test No. 316: Phototransformation of Chemicals in Water - Direct Photolysis*; Organisation for Economic Co-operation and Development: Paris, 2008.

(10) OECD *Test No. 306: Biodegradability in Seawater*; Organisation for Economic Co-operation and Development: Paris, 1992.

(11) OECD *Test No. 309: Aerobic Mineralisation in Surface Water - Simulation Biodegradation Test*; Organisation for Economic Co-operation and Development: Paris, 2004.

(12) OECD *Test No. 308: Aerobic and Anaerobic Transformation in Aquatic Sediment Systems*; Organisation for Economic Co-operation and Development: Paris, 2002.

(13) Lai, F. Y.; Erratico, C.; Kinyua, J.; Mueller, J. F.; Covaci, A.; van Nuijs, A. L. N. Liquid Chromatography-Quadrupole Time-of-Flight Mass Spectrometry for Screening in Vitro Drug Metabolites in Humans: Investigation on Seven Phenethylamine-Based Designer Drugs. *J. Pharm. Biomed. Anal.* **2015**, *114*, 355−375.

(14) Baduel, C.; Lai, F. Y.; van Nuijs, A. L. N.; Covaci, A. Suspect and Nontargeted Strategies to Investigate *in Vitro* Human Biotransformation Products of Emerging Environmental Contaminants: The Benzotriazoles. *Environ. Sci. Technol.* **2019**, *53* (17), 10462−10469.

(15) Van Eerd, L. L.; Hoagland, R. E.; Zablotowicz, R. M.; Hall, J. C. Pesticide Metabolism in Plants and Microorganisms. *Weed Sci.* **2003**, *51* (4), 472−495.

(16) Desiante, W. L.; Carles, L.; Wullschleger, S.; Joss, A.; Stamm, C.; Fenner, K. Wastewater Microorganisms Impact the Micropollutant Biotransformation Potential of Natural Stream Biofilms. *Water Res.* **2022**, *217*, 118413.

(17) Olvera-Vargas, H.; Leroy, S.; Rivard, M.; Oturan, N.; Oturan, M.; Buisson, D. Microbial Biotransformation of Furosemide for Environmental Risk Assessment: Identification of Metabolites and Toxicological Evaluation. *Environ. Sci. Pollut. Res.* **2016**, *23* (22), 22691−22700.

(18) Marinho, A. T.; Rodrigues, P. M.; Caixas, U.; Antunes, A. M. M.; Branco, T.; Harjivan, S. G.; Marques, M. M.; Monteiro, E. C.; Pereira, S. A. Differences in Nevirapine Biotransformation as a Factor for Its Sex-Dependent Dimorphic Profile of Adverse Drug Reactions. *J. Antimicrob. Chemother.* **2014**, *69* (2), 476−482.

(19) Choi, Y.; Jeon, J.; Kim, S. D. Identification of Biotransformation Products of Organophosphate Ester from Various Aquatic Species by Suspect and Non-Target Screening Approach. *Water Res.* **2021**, *200*, 117201.

(20) Löffler, P.; Henschel, H.; Ugolini, V.; Flores Quintana, H.; Wiberg, K.; Lai, F. Y. Exploring the Role of Photolysis in the Aquatic Fate of Antimicrobial Transformation Products: Implications for One Health. *ACS EST Water* **2025**, *5* (7), 4112−4119.

(21) Helmus, R.; Bagdonaite, I.; de Voogt, P.; van Bommel, M. R.; Schymanski, E. L.; van Wezel, A. P.; ter Laak, T. L. Comprehensive Mass Spectrometry Workflows to Systematically Elucidate Transformation Processes of Organic Micropollutants: A Case Study on the Photodegradation of Four Pharmaceuticals. *Environ. Sci. Technol.* **2025**, *59* (7), 3723−3736.

(22) Bourgin, M.; Borowska, E.; Helbing, J.; Hollender, J.; Kaiser, H.-P.; Kienle, C.; McArdell, C. S.; Simon, E.; von Gunten, U. Effect of Operational and Water Quality Parameters on Conventional Ozonation and the Advanced Oxidation Process O3/H2O2: Kinetics of Micropollutant Abatement, Transformation Product and Bromate Formation in a Surface Water. *Water Res.* **2017**, *122*, 234−245.

(23) Ooi, G. T. H.; Escola Casas, M.; Andersen, H. R.; Bester, K. Transformation Products of Clindamycin in Moving Bed Biofilm Reactor (MBBR). *Water Res.* **2017**, *113*, 139−148.

(24) Adamek, E.; Baran, W.; Sobczak, A. Photocatalytic Degradation of Veterinary Antibiotics: Biodegradability and Antimicrobial Activity of Intermediates. *Process Saf. Environ. Prot.* **2016**, *103*, 1−9.

(25) Adamek, E.; Baran, W. Degradation of Veterinary Antibiotics by the Ozonation Process: Product Identification and Ecotoxicity Assessment. *J. Hazard. Mater.* **2024**, *469*, 134026.

(26) Richardson, S. D.; Manasfi, T. Water Analysis: Emerging Contaminants and Current Issues. *Anal. Chem.* **2024**, *96* (20), 8184−8219.

(27) Bletsou, A. A.; Jeon, J.; Hollender, J.; Archontaki, E.; Thomaidis, N. S. Targeted and Non-Targeted Liquid Chromatography-Mass Spectrometric Workflows for Identification of Transformation Products of Emerging Pollutants in the Aquatic Environment. *TrAC, Trends Anal. Chem.* **2015**, *66*, 32−44.

(28) Hollender, J.; Schymanski, E. L.; Ahrens, L.; Alygizakis, N.; Béen, F.; Bijlsma, L.; Brunner, A. M.; Celma, A.; Fildier, A.; Fu, Q.; Gago-Ferrero, P.; Gil-Solsona, R.; Haglund, P.; Hansen, M.; Kaserzon, S.; Kruve, A.; Lamoree, M.; Margoum, C.; Meijer, J.; Merel, S.; Rauert, C.; Rostkowski, P.; Samanipour, S.; Schulze, B.; Schulze, T.; Singh, R. R.; Slobodnik, J.; Steininger-Mairinger, T.; Thomaidis, N. S.; Togola, A.; Vorkamp, K.; Vulliet, E.; Zhu, L.; Krauss, M. NORMAN Guidance on Suspect and Non-Target Screening in Environmental Monitoring. *Environ. Sci. Eur.* **2023**, *35* (1), 75.

(29) Mercurio, P.; Eaglesham, G.; Parks, S.; Kenway, M.; Beltran, V.; Flores, F.; Mueller, J. F.; Negri, A. P. Contribution of Transformation Products towards the Total Herbicide Toxicity to Tropical Marine Organisms. *Sci. Rep.* **2018**, *8* (1), 4808.

(30) Tian, Z.; Gonzalez, M.; Rideout, C. A.; Zhao, H. N.; Hu, X.; Wetzel, J.; Mudrock, E.; James, C. A.; McIntyre, J. K.; Kolodziej, E. P. 6PPD-Quinone: Revised Toxicity Assessment and Quantification with a Commercial Standard. *Environ. Sci. Technol. Lett.* **2022**, *9* (2), 140−146.

(31) Jin, Y.; Wang, L.; Chen, G.; Lin, X.; Miao, W.; Fu, Z. Exposure of Mice to Atrazine and Its Metabolite Diaminochlorotriazine Elicits Oxidative Stress and Endocrine Disruption. *Environ. Toxicol. Pharmacol.* **2014**, *37* (2), 782−790.

(32) Bender, R. P.; Lindsey, R. H.; Burden, D. A.; Osheroff, N. N-Acetyl-p-Benzoquinone Imine, the Toxic Metabolite of Acetamino-

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

phen, Is a Topoisomerase II Poison. *Biochemistry* **2004**, *43* (12), 3731−3739.
(33) Wetterauer, B.; Ricking, M.; Otte, J. C.; Hallare, A. V.; Rastall, A.; Erdinger, L.; Schwarzbauer, J.; Braunbeck, T.; Hollert, H. Toxicity, Dioxin-like Activities, and Endocrine Effects of DDT Metabolites—DDA; DDMU, DDMS, and DDCN. *Environ. Sci. Pollut. Res.* **2012**, *19* (2), 403−415.
(34) Plowchalk, D. R.; Mattison, D. R. Phosphoramide Mustard Is Responsible for the Ovarian Toxicity of Cyclophosphamide. *Toxicol. Appl. Pharmacol.* **1991**, *107* (3), 472−481.
(35) Maculewicz, J.; Kowalska, D.; Świacka, K.; Toński, M.; Stepnowski, P.; Białk-Bielińska, A.; Dołżonek, J. Transformation Products of Pharmaceuticals in the Environment: Their Fate, (Eco)Toxicity and Bioaccumulation Potential. *Sci. Total Environ.* **2022**, *802*, 149916.
(36) Ji, C.; Song, Q.; Chen, Y.; Zhou, Z.; Wang, P.; Liu, J.; Sun, Z.; Zhao, M. The Potential Endocrine Disruption of Pesticide Transformation Products (TPs): The Blind Spot of Pesticide Risk Assessment. *Environ. Int.* **2020**, *137*, 105490.
(37) Neuwoehner, J.; Fenner, K.; Escher, B. I. Physiological Modes of Action of Fluoxetine and Its Human Metabolites in Algae. *Environ. Sci. Technol.* **2009**, *43* (17), 6830−6837.
(38) Nałęcz-Jawecki, G.; Wójcik, T.; Sawicki, J. Evaluation of in Vitro Biotransformation of Propranolol with HPLC, MS/MS, and Two Bioassays. *Environ. Toxicol.* **2008**, *23* (1), 52−58.
(39) Tian, Z.; Zhao, H.; Peter, K. T.; Gonzalez, M.; Wetzel, J.; Wu, C.; Hu, X.; Prat, J.; Mudrock, E.; Hettinger, R.; Cortina, A. E.; Biswas, R. G.; Kock, F. V. C.; Soong, R.; Jenne, A.; Du, B.; Hou, F.; He, H.; Lundeen, R.; Gilbreath, A.; Sutton, R.; Scholz, N. L.; Davis, J. W.; Dodd, M. C.; Simpson, A.; McIntyre, J. K.; Kolodziej, E. P. A Ubiquitous Tire Rubber-Derived Chemical Induces Acute Mortality in Coho Salmon. *Science* **2021**, *371* (6525), 185−189.
(40) Tian, Z.; Zhao, H.; Peter, K. T.; Gonzalez, M.; Wetzel, J.; Wu, C.; Hu, X.; Prat, J.; Mudrock, E.; Hettinger, R.; Cortina, A. E.; Biswas, R. G.; Kock, F. V. C.; Soong, R.; Jenne, A.; Du, B.; Hou, F.; He, H.; Lundeen, R.; Gilbreath, A.; Sutton, R.; Scholz, N. L.; Davis, J. W.; Dodd, M. C.; Simpson, A.; McIntyre, J. K.; Kolodziej, E. P. Erratum for the Report "A Ubiquitous Tire Rubber-Derived Chemical Induces Acute Mortality in Coho Salmon,.". *Science* **2022**, *375* (6582), No. eabo5785.
(41) Chen, X.; He, T.; Yang, X.; Gan, Y.; Qing, X.; Wang, J.; Huang, Y. Analysis, Environmental Occurrence, Fate and Potential Toxicity of Tire Wear Compounds 6PPD and 6PPD-Quinone. *J. Hazard. Mater.* **2023**, *452*, 131245.
(42) Brinkmann, M.; Montgomery, D.; Selinger, S.; Miller, J. G. P.; Stock, E.; Alcaraz, A. J.; Challis, J. K.; Weber, L.; Janz, D.; Hecker, M.; Wiseman, S. Acute Toxicity of the Tire Rubber-Derived Chemical 6PPD-Quinone to Four Fishes of Commercial, Cultural, and Ecological Importance. *Environ. Sci. Technol. Lett.* **2022**, *9* (4), 333−338.
(43) Cao, G.; Wang, W.; Zhang, J.; Wu, P.; Zhao, X.; Yang, Z.; Hu, D.; Cai, Z. New Evidence of Rubber-Derived Quinones in Water, Air, and Soil. *Environ. Sci. Technol.* **2022**, *56* (7), 4142−4150.
(44) Escher, B. I.; Stapleton, H. M.; Schymanski, E. L. Tracking Complex Mixtures of Chemicals in Our Changing Environment. *Science* **2020**, *367* (6476), 388−392.
(45) Kamerlin, N.; Delcey, M. G.; Manzetti, S.; van der Spoel, D. Toward a Computational Ecotoxicity Assay. *J. Chem. Inf. Model.* **2020**, *60* (8), 3792−3803.
(46) Löffler, P.; Lai, A.; Henschel, H.; Spilsbury, F.; Deviller, G.; Tarazona, J. V.; Lai, F. Y. Enhanced Risk Assessment of Transformation Products through Chemical Similarity Analysis. *ACS EST Water* **2024**, *4* (5), 1949−1951.
(47) European Chemicals Agency (ECHA) *Guidance on Information Requirements and Chemical Safety Assessment - Chapter R.6: QSARs and Grouping of Chemicals*, 2008.
(48) European Chemicals Agency *Read-Across Assessment Framework (RAAF)*; Publications Office: LU, 2017.
(49) Zahn, D.; Arp, H. P. H.; Fenner, K.; Georgi, A.; Hafner, J.; Hale, S. E.; Hollender, J.; Letzel, T.; Schymanski, E. L.; Sigmund, G.; Reemtsma, T. Should Transformation Products Change the Way We Manage Chemicals? *Environ. Sci. Technol.* **2024**, *58* (18), 7710−7718.
(50) Halldin, K.; Berg, C.; Larsson, L.; Leto, E. *Chemical Pollution and One Health - from Reactivity to Proactivity: Conflicting Objectives - Using Effective Drugs without Polluting Our Environment*; Uppsala, 2023.
(51) Wiberg, K.; Lai, F. Y.; Ahrens, L.; Ugolini, V.; Löffler, P.; Arnell, M.; Lavonen, E.; Arp, H. P. H. *Chemical Pollution and One Health - from Reactivity to Proactivity: Water Quality in One Health: Managing Chemical Risks*; Uppsala, 2023; Vol. *2*, pp 482−509.
(52) Ashby, J.; Tennant, R. W. Definitive Relationships among Chemical Structure, Carcinogenicity and Mutagenicity for 301 Chemicals Tested by the U.S. NTP. *Mutat. Res. Rev. Genet. Toxicol.* **1991**, *257* (3), 229−306.
(53) Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; Consonni, V.; Kuz'min, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, *57* (12), 4977−5010.
(54) Tate, T.; Wambaugh, J.; Patlewicz, G.; Shah, I. Repeat-Dose Toxicity Prediction with Generalized Read-Across (GenRA) Using Targeted Transcriptomic Data: A Proof-of-Concept Case Study. *Comput. Toxicol.* **2021**, *19*, 100171.
(55) Wu, J.; D'Ambrosi, S.; Ammann, L.; Stadnicka-Michalak, J.; Schirmer, K.; Baity-Jesi, M. Predicting Chemical Hazard across Taxa through Machine Learning. *Environ. Int.* **2022**, *163*, 107184.
(56) van de Schoot, R.; de Bruin, J.; Schram, R.; Zahedi, P.; de Boer, J.; Weijdema, F.; Kramer, B.; Huijts, M.; Hoogerwerf, M.; Ferdinands, G.; Harkema, A.; Willemsen, J.; Ma, Y.; Fang, Q.; Hindriks, S.; Tummers, L.; Oberski, D. L. An Open Source Machine Learning Framework for Efficient and Transparent Systematic Reviews. *Nat. Mach. Intell.* **2021**, *3* (2), 125−133.
(57) Kebede, M. M.; Le Cornet, C.; Fortner, R. T. In-Depth Evaluation of Machine Learning Methods for Semi-Automating Article Screening in a Systematic Review of Mechanistic Literature. *Res. Synth. Methods* **2023**, *14* (2), 156−172.
(58) Pham, B.; Jovanovic, J.; Bagheri, E.; Antony, J.; Ashoor, H.; Nguyen, T. T.; Rios, P.; Robson, R.; Thomas, S. M.; Watt, J.; Straus, S. E.; Tricco, A. C. Text Mining to Support Abstract Screening for Knowledge Syntheses: A Semi-Automated Workflow. *Syst. Rev.* **2021**, *10* (1), 156.
(59) Rayyan - Intelligent Systematic Review - Rayyan. https://www.rayyan.ai/(accessed March 18, 2024).
(60) Swain, M. C.; Cole, J. M. ChemDataExtractor: A Toolkit for Automated Extraction of Chemical Information from the Scientific Literature. *J. Chem. Inf. Model.* **2016**, *56* (10), 1894−1904.
(61) Dong, Q.; Cole, J. M. Snowball 2.0: Generic Material Data Parser for ChemDataExtractor. *J. Chem. Inf. Model.* **2023**, *63* (22), 7045−7055.
(62) Palm, E. H.; Chirsir, P.; Krier, J.; Thiessen, P. A.; Zhang, J.; Bolton, E. E.; Schymanski, E. L. ShinyTPs: Curating Transformation Products from Text Mining Results. *Environ. Sci. Technol. Lett.* **2023**, *10* (10), 865−871.
(63) Wicker, J.; Lorsbach, T.; Gütlein, M.; Schmid, E.; Latino, D.; Kramer, S.; Fenner, K. EnviPath - The Environmental Contaminant Biotransformation Pathway Resource. *Nucleic Acids Res.* **2016**, *44* (D1), D502−D508.
(64) Hafner, J.; Lorsbach, T.; Schmidt, S.; Brydon, L.; Dost, K.; Zhang, K.; Fenner, K.; Wicker, J. Advancements in Biotransformation Pathway Prediction: Enhancements, Datasets, and Novel Functionalities in EnviPath. *Journal of Chem. Inf.* **2024**, *16* (1), 93.
(65) Mohammed Taha, H.; Aalizadeh, R.; Alygizakis, N.; Antignac, J.-P.; Arp, H. P. H.; Bade, R.; Baker, N.; Belova, L.; Bijlsma, L.; Bolton, E. E.; Brack, W.; Celma, A.; Chen, W.-L.; Cheng, T.; Chirsir, P.; Čirka, Ľ.; D'Agostino, L. A.; Djoumbou Feunang, Y.; Dulio, V.; Fischer, S.; Gago-Ferrero, P.; Galani, A.; Geueke, B.; Głowacka, N.;

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Glüge, J.; Groh, K.; Grosse, S.; Haglund, P.; Hakkinen, P. J.; Hale, S. E.; Hernandez, F.; Janssen, E. M.-L.; Jonkers, T.; Kiefer, K.; Kirchner, M.; Koschorreck, J.; Krauss, M.; Krier, J.; Lamoree, M. H.; Letzel, M.; Letzel, T.; Li, Q.; Little, J.; Liu, Y.; Lunderberg, D. M.; Martin, J. W.; McEachran, A. D.; McLean, J. A.; Meier, C.; Meijer, J.; Menger, F.; Merino, C.; Muncke, J.; Muschket, M.; Neumann, M.; Neveu, V.; Ng, K.; Oberacher, H.; O'Brien, J.; Oswald, P.; Oswaldova, M.; Picache, J. A.; Postigo, C.; Ramirez, N.; Reemtsma, T.; Renaud, J.; Rostkowski, P.; Rüdel, H.; Salek, R. M.; Samanipour, S.; Scheringer, M.; Schliebner, I.; Schulz, W.; Schulze, T.; Sengl, M.; Shoemaker, B. A.; Sims, K.; Singer, H.; Singh, R. R.; Sumarah, M.; Thiessen, P. A.; Thomas, K. V.; Torres, S.; Trier, X.; van Wezel, A. P.; Vermeulen, R. C. H.; Vlaanderen, J. J.; von der Ohe, P. C.; Wang, Z.; Williams, A. J.; Willighagen, E. L.; Wishart, D. S.; Zhang, J.; Thomaidis, N. S.; Hollender, J.; Slobodnik, J.; Schymanski, E. L. The NORMAN Suspect List Exchange (NORMAN-SLE): Facilitating European and Worldwide Collaboration on Suspect Screening in High Resolution Mass Spectrometry. *Environ. Sci. Eur.* **2022**, *34* (1), 104.

(66) Schymanski, E.; Bolton, E. E.; Cheng, T.; Thiessen, P. A.; Zhang, J.; Helmus, R.; Blanke, G. *Transformations in PubChem - Full Dataset*, 2025.

(67) Alygizakis, N.; Jonkers, T. J. H. *S6 | ITNANTIBIOTIC | Antibiotic List: ITN MSCA ANSWER*, 2022.

(68) Renaud, J.; Sumarah, M. *S26 | MYCOTOXINS | List of Mycotoxins from AAFC*, 2023.

(69) LMC, Slobodnik, J.. *S38 | SOLNSLMCTPS | SOLUTIONS Predicted Transformation Products by LMC*, 2019.

(70) Kiefer, K.; Müller, A.; Singer, H.; Hollender, J. *S60 | SWISSPEST19 | Swiss Pesticides and Metabolites from Kiefer et al 2019*, 2023.

(71) Schollee, J.; Schymanski, E.; Stravs, M.; Gulde, R.; Thomaidis, N.; Hollender, J. *S66 | EAWAGTPS | Parent-Transformation Product Pairs from Eawag*, 2024.

(72) LCSB-ECI; Krier, J.; Palm, E.; Chirsir, P.; Komolo, J.; Lontsie Zanmene, M.; Schymanski, E.; Bolton, E. E.; Thiessen, P. A.; Zhang, J.; PubChem Team. *S68 | HSDBTPS | Transformation Products Extracted from HSDB Content in PubChem*, 2023.

(73) Meijer, J.; Lamoree, M.; Hamers, T.; Antingac, J.-P.; Hutinet, S.; Debrauwer, L.; Covaci, A.; Huber, C.; Krauss, M.; Walker, D. I.; Schymanski, E.; Vermeulen, R.; Vlaanderen, J. J. *S71 | CECSCREEN | HBM4 EU CECscreen: Screening List for Chemicals of Emerging Concern Plus Metadata and Predicted Phase 1 Metabolites*, 2022.

(74) Djoumbou-Feunang, Y.; Schymanski, E.; Zhang, J.; Wishart, D. S. *S73 | METXBIODB | Metabolite Reaction Database from BioTransformer*, 2024.

(75) Schymanski, E. L.; Baesu, A.; Chirsir, P.; Löffler, P.; Helmus, R. *S74 | REFTPS | Transformation Products and Reactions from Literature*, 2024.

(76) Janssen, E. M.-L.; Jones, M. R.; Pinto, E.; Dörr, F.; Torres, M. A.; Rios Jacinavicius, F.; Mazur-Marzec, H.; Szubert, K.; Konkel, R.; Tartaglione, L.; Dell'Aversano, C.; Miglione, A.; McCarron, P.; Beach, D. G.; Miles, C. O.; Fewer, D. P.; Sivonen, K.; Jokela, J.; Wahlsten, M.; Niedermeyer, T. H. J.; Schanbacher, F.; Leão, P.; Preto, M.; D'Agostino, P. M.; Baunach, M.; Dittmann, E.; Reher, R. *S75 | CyanoMetDB | Comprehensive Database of Secondary Metabolites from Cyanobacteria*, 2023.

(77) Menger, F.; Boström, G. *S78 | SLUPESTTPS | Pesticides and TPs from SLU, Sweden (NORMAN-SLE-S78.0.1.3) [Data Set]*, 2023.

(78) Belova, L.; Caballero-Casero, N.; van Nuijs, A. L. N.; Covaci, A. *S79 | UACCSCEC | Collision Cross Section (CCS) Library from UAntwerp*, 2022.

(79) Merino, C.; Vinaixa, M.; Ramirez, N. *S81 | THSTPS | Thirdhand Smoke Specific Metabolites*, 2021.

(80) Umweltbundesamt, Mohammed Taha, H.. *S88 | UBABIOCIDES | List of Prioritized Biocides from UBA*, 2021.

(81) Meyer, C.; Hollender, J. *S113 | SWISSPHARMA24 | 2024 Swiss Pharmaceutical List with Metabolites*, 2024.

(82) Löffler, P.; Lai, F. Y. *S114 | SLUAMTPS | Antimicrobial Transformation Products from SLU (NORMAN-SLE-S114.0.1.0) [Data Set]*, 2024.

(83) Schymanski, E. L.; Bolton, E. E. FAIRifying the Exposome Journal: Templates for Chemical Structures and Transformations. *Exposome* **2022**, *2* (1), osab006.

(84) Helmus, R.; ter Laak, T. L.; van Wezel, A. P.; de Voogt, P.; Schymanski, E. L. PatRoon: Open Source Software Platform for Environmental Mass Spectrometry Based Non-Target Screening. *J. Chem. Inf.* **2021**, *13* (1), 1.

(85) Helmus, R.; van de VeldeBrunner, B. A. M.; ter Laak, T. L.; van Wezel, A. P.; Schymanski, E. L.; Schymanski, E. L. PatRoon 2.0: Improved Non-Target Analysis Workflows Including Automated Transformation Product Screening. *J. Open Source Softw.* **2022**, *7* (71), 4029.

(86) Djoumbou-Feunang, Y.; Fiamoncini, J.; Gil-de-la-Fuente, A.; Greiner, R.; Manach, C.; Wishart, D. S. BioTransformer: A Comprehensive Computational Tool for Small Molecule Metabolism Prediction and Metabolite Identification. *Journal of Chem. Inf.* **2019**, *11* (1), 2.

(87) Wishart, D. S.; Tian, S.; Allen, D.; Oler, E.; Peters, H.; Lui, V. W.; Gautam, V.; Djoumbou-Feunang, Y.; Greiner, R.; Metz, T. O. BioTransformer 3.0—a Web Server for Accurately Predicting Metabolic Transformation Products. *Nucleic Acids Res.* **2022**, *50* (W1), W115−W123.

(88) Zhu, J.-J.; Jiang, J.; Yang, M.; Ren, Z. J. ChatGPT and Environmental Research. *Environ. Sci. Technol.* **2023**, *57* (46), 17667−17670.

(89) Hicks, M. T.; Humphries, J.; Slater, J. ChatGPT Is Bullshit. *Ethics Inf. Technol.* **2024**, *26* (2), 38.

(90) Zimmerman, J. B. Author Guidelines Environmental Science & Technology. Author Guidelines Environmental Science & Technology. https://researcher-resources.acs.org/publish/author_guidelines?coden=esthag#data_requirements (accessed July 23, 2025).

(91) Djoumbou Founang, Y. djoumbou/biotransformerjar/database — Bitbucket. Bitbucket. https://bitbucket.org/djoumbou/biotransformerjar/src/master/database/(accessed Sep 06, 2024).

(92) Wolfe, K.; Pope, N.; Parmar, R.; Galvin, M.; Stevens, C.; Weber, E.; Flaishans, J.; Purucker, T. *Chemical Transformation System: Cloud Based Cheminformatic Services to Support Integrated Environmental Modeling*; International Congress on Environmental Modelling and Software, 2016.

(93) Ellis, L. B. M.; Wackett, L. P. A Microbial Biocatalysis Database. *Soc. Ind. Microbiol. News* **1995**, *45* (4), 167−173.

(94) Gao, J.; Ellis, L. B. M.; Wackett, L. P. The University of Minnesota Biocatalysis/Biodegradation Database: Improving Public Access. *Nucleic Acids Res.* **2010**, *38* (suppl_1), D488−D491.

(95) Kanehisa, M.; Furumichi, M.; Sato, Y.; Matsuura, Y.; Ishiguro-Watanabe, M. KEGG: Biological Systems Database as a Model of the Real World. *Nucleic Acids Res.* **2025**, *53* (D1), D672−D677.

(96) Rich, S.; Higgins, C.; Helbling, D. E.; Fenner, K. EnviPath-PFAS: A Publicly Available, Expert-Curated Database and Pathway Prediction System for Biotransformations of Precursors to Persistent Perfluorinated Alkyl Substances (PFASs). In *SETAC Europe 35th Annual Meeting: SETAC Europe 35th Annual Meeting*, 2025.

(97) US EPA *Estimation Programs Interface Suite for Microsoft Windows*; United States Environmental Protection Agency: Washington, DC, USA, 2012.

(98) Rodríguez-Leal, I.; MacLeod, M. The Applicability Domain of EPI Suite™ for Screening Phytotoxins for Potential to Contaminate Source Water for Drinking. *Environ. Sci. Eur.* **2022**, *34* (1), 96.

(99) Zhang, Z.; Sangion, A.; Wang, S.; Gouin, T.; Brown, T.; Arnot, J. A.; Li, L. Chemical Space Covered by Applicability Domains of Quantitative Structure-Property Relationships and Semiempirical Relationships in Chemical Assessments. *Environ. Sci. Technol.* **2024**, *58* (7), 3386−3398.

(100) Patlewicz, G.; Jeliazkova, N.; Safford, R. J.; Worth, A. P.; Aleksiev, B. An Evaluation of the Implementation of the Cramer

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

Classification Scheme in the Toxtree Software. *SAR QSAR Environ. Res.* **2008**, *19* (5−6), 495−524.
(101) Martin, T. M. *User's Guide for T.E.S.T. (Toxicity Estimation Software Tool)*; U.S. Environmental Protection Agency, 2020.
(102) Mansouri, K.; Grulke, C. M.; Judson, R. S.; Williams, A. J. OPERA Models for Predicting Physicochemical Properties and Environmental Fate Endpoints. *J. Chem. Inf.* **2018**, *10* (1), 10.
(103) Benfenati, E.; Manganaro, A.; Gini, G. VEGA-QSAR: AI inside a Platform for Predictive Toxicology; *CEUR Workshop Proceedings* **2013**; Vol. *1107*, pp 21−28.
(104) Gustavsson, M.; Käll, S.; Svedberg, P.; Inda-Diaz, J. S.; Molander, S.; Coria, J.; Backhaus, T.; Kristiansson, E. Transformers Enable Accurate Prediction of Acute and Chronic Chemical Toxicity in Aquatic Organisms. *Sci. Adv.* **2024**, *10* (10), No. eadk6669.
(105) Ramaprasad, A. S. E.; Smith, M. T.; McCoy, D.; Hubbard, A. E.; La Merrill, M. A.; Durkin, K. A. Predicting the Binding of Small Molecules to Nuclear Receptors Using Machine Learning. *Briefings Bioinf.* **2022**, *23* (3), bbac114.
(106) Zhu, H.; Martin, T. M.; Ye, L.; Sedykh, A.; Young, D. M.; Tropsha, A. Quantitative Structure-Activity Relationship Modeling of Rat Acute Toxicity by Oral Exposure. *Chem. Res. Toxicol.* **2009**, *22* (12), 1913−1921.
(107) Martin, T. M.; Young, D. M. Prediction of the Acute Toxicity (96-h LC50) of Organic Compounds to the Fathead Minnow (Pimephales Promelas) Using a Group Contribution Method. *Chem. Res. Toxicol.* **2001**, *14* (10), 1378−1385.
(108) Martin, T. M.; Harten, P.; Venkatapathy, R.; Das, S.; Young, D. M. A Hierarchical Clustering Methodology for the Estimation of Toxicity. *Toxicol. Mech. Methods* **2008**, *18* (2−3), 251−266.
(109) Young, D.; Martin, T.; Venkatapathy, R.; Harten, P. Are the Chemical Structures in Your QSAR Correct? *QSAR Comb. Sci.* **2008**, *27* (11−12), 1337−1345.
(110) Sushko, I.; Novotarskyi, S.; Körner, R.; Pandey, A. K.; Cherkasov, A.; Li, J.; Gramatica, P.; Hansen, K.; Schroeter, T.; Müller, K.-R.; Xi, L.; Liu, H.; Yao, X.; Öberg, T.; Hormozdiari, F.; Dao, P.; Sahinalp, C.; Todeschini, R.; Polishchuk, P.; Artemenko, A.; Kuz'min, V.; Martin, T. M.; Young, D. M.; Fourches, D.; Muratov, E.; Tropsha, A.; Baskin, I.; Horvath, D.; Marcou, G.; Muller, C.; Varnek, A.; Prokopenko, V. V.; Tetko, I. V. Applicability Domains for Classification Problems: Benchmarking of Distance to Models for Ames Mutagenicity Set. *J. Chem. Inf. Model.* **2010**, *50* (12), 2094−2111.
(111) Mansouri, K.; Abdelaziz, A.; Rybacka, A.; Roncaglioni, A.; Tropsha, A.; Varnek, A.; Zakharov, A.; Worth, A.; Richard, A. M.; Grulke, C. M.; Trisciuzzi, D.; Fourches, D.; Horvath, D.; Benfenati, E.; Muratov, E.; Wedebye, E. B.; Grisoni, F.; Mangiatordi, G. F.; Incisivo, G. M.; Hong, H.; Ng, H. W.; Tetko, I. V.; Balabin, I.; Kancherla, J.; Shen, J.; Burton, J.; Nicklaus, M.; Cassotti, M.; Nikolov, N. G.; Nicolotti, O.; Andersson, P. L.; Zang, Q.; Politi, R.; Beger, R. D.; Todeschini, R.; Huang, R.; Farag, S.; Rosenberg, S. A.; Slavov, S.; Hu, X.; Judson, R. S. CERAPP: Collaborative Estrogen Receptor Activity Prediction Project. *Environ. Health Perspect.* **2016**, *124* (7), 1023−1033.
(112) Mansouri, K.; Kleinstreuer, N.; Abdelaziz, A. M.; Alberga, D.; Alves, V. M.; Andersson, P. L.; Andrade, C. H.; Bai, F.; Balabin, I.; Ballabio, D.; Benfenati, E.; Bhhatarai, B.; Boyer, S.; Chen, J.; Consonni, V.; Farag, S.; Fourches, D.; García-Sosa, A. T.; Gramatica, P.; Grisoni, F.; Grulke, C. M.; Hong, H.; Horvath, D.; Hu, X.; Huang, R.; Jeliazkova, N.; Li, J.; Li, X.; Liu, H.; Manganelli, S.; Mangiatordi, G. F.; Maran, U.; Marcou, G.; Martin, T.; Muratov, E.; Nguyen, D.-T.; Nicolotti, O.; Nikolov, N. G.; Norinder, U.; Papa, E.; Petitjean, M.; Piir, G.; Pogodin, P.; Poroikov, V.; Qiao, X.; Richard, A. M.; Roncaglioni, A.; Ruiz, P.; Rupakheti, C.; Sakkiah, S.; Sangion, A.; Schramm, K.-W.; Selvaraj, C.; Shah, I.; Sild, S.; Sun, L.; Taboureau, O.; Tang, Y.; Tetko, I. V.; Todeschini, R.; Tong, W.; Trisciuzzi, D.; Tropsha, A.; Van Den Driessche, G.; Varnek, A.; Wang, Z.; Wedebye, E. B.; Williams, A. J.; Xie, H.; Zakharov, A. V.; Zheng, Z.; Judson, R. S. CoMPARA: Collaborative Modeling Project for Androgen Receptor Activity. *Environ. Health Perspect.* **2020**, *128* (2), 027002.
(113) Mansouri, K.; Karmaus, A. L.; Fitzpatrick, J.; Patlewicz, G.; Pradeep, P.; Alberga, D.; Alepee, N.; Allen, T. E. H.; Allen, D.; Alves, V. M.; Andrade, C. H.; Auernhammer, T. R.; Ballabio, D.; Bell, S.; Benfenati, E.; Bhattacharya, S.; Bastos, J. V.; Boyd, S.; Brown, J. B.; Capuzzi, S. J.; Chushak, Y.; Ciallella, H.; Clark, A. M.; Consonni, V.; Daga, P. R.; Ekins, S.; Farag, S.; Fedorov, M.; Fourches, D.; Gadaleta, D.; Gao, F.; Gearhart, J. M.; Goh, G.; Goodman, J. M.; Grisoni, F.; Grulke, C. M.; Hartung, T.; Hirn, M.; Karpov, P.; Korotcov, A.; Lavado, G. J.; Lawless, M.; Li, X.; Luechtefeld, T.; Lunghini, F.; Mangiatordi, G. F.; Marcou, G.; Marsh, D.; Martin, T.; Mauri, A.; Muratov, E. N.; Myatt, G. J.; Nguyen, D.-T.; Nicolotti, O.; Note, R.; Pande, P.; Parks, A. K.; Peryea, T.; Polash, A. H.; Rallo, R.; Roncaglioni, A.; Rowlands, C.; Ruiz, P.; Russo, D. P.; Sayed, A.; Sayre, R.; Sheils, T.; Siegel, C.; Silva, A. C.; Simeonov, A.; Sosnin, S.; Southall, N.; Strickland, J.; Tang, Y.; Teppen, B.; Tetko, I. V.; Thomas, D.; Tkachenko, V.; Todeschini, R.; Toma, C.; Tripodi, I.; Trisciuzzi, D.; Tropsha, A.; Varnek, A.; Vukovic, K.; Wang, Z.; Wang, L.; Waters, K. M.; Wedlake, A. J.; Wijeyesakere, S. J.; Wilson, D.; Xiao, Z.; Yang, H.; Zahoranszky-Kohalmi, G.; Zakharov, A. V.; Zhang, F. F.; Zhang, Z.; Zhao, T.; Zhu, H.; Zorn, K. M.; Casey, W.; Kleinstreuer, N. C. CATMoS: Collaborative Acute Toxicity Modeling Suite. *Environ. Health Perspect.* **2021**, *129* (4), 047013.
(114) Mansouri, K.; Cariello, N. F.; Korotcov, A.; Tkachenko, V.; Grulke, C. M.; Sprankle, C. S.; Allen, D.; Casey, W. M.; Kleinstreuer, N. C.; Williams, A. J. Open-Source QSAR Models for PKa Prediction Using Multiple Machine Learning Approaches. *J. Chem. Inf.* **2019**, *11* (1), 60.
(115) Willighagen, E. L.; Mayfield, J. W.; Alvarsson, J.; Berg, A.; Carlsson, L.; Jeliazkova, N.; Kuhn, S.; Pluskal, T.; Rojas-Chertó, M.; Spjuth, O.; Torrance, G.; Evelo, C. T.; Guha, R.; Steinbeck, C. The Chemistry Development Kit (CDK) v2.0: Atom Typing, Depiction, Molecular Formulas, and Substructure Searching. *J. Chem. Inf.* **2017**, *9* (1), 33.
(116) Yap, C. W. PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints. *J. Comput. Chem.* **2011**, *32* (7), 1466−1474.
(117) 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.; Richard, A. M. The CompTox Chemistry Dashboard: A Community Data Resource for Environmental Chemistry. *J. Chem. Inf.* **2017**, *9* (1), 61.
(118) OECD. *The OECD QSAR Toolbox - Developed by the OECD and the European Chemicals Agency (ECHA)*, 2025.
(119) Schultz, T. W.; Diderich, R.; Kuseva, C. D.; Mekenyan, O. G. The OECD QSAR Toolbox Starts Its Second Decade. In *Computational Toxicology*; Nicolotti, O., Ed.; Methods in Molecular Biology; Springer New York: New York, NY, 2018; Vol. *1800*, pp 55−77..
(120) Bishop, P. L.; Mansouri, K.; Eckel, W. P.; Lowit, M. B.; Allen, D.; Blankinship, A.; Lowit, A. B.; Harwood, D. E.; Johnson, T.; Kleinstreuer, N. C. Evaluation of in Silico Model Predictions for Mammalian Acute Oral Toxicity and Regulatory Application in Pesticide Hazard and Risk Assessment. *Regul. Toxicol. Pharmacol.* **2024**, *149*, 105614.
(121) Benfenati, E. The CAESAR Project for in Silico Models for the REACH Legislation. *Chem. Cent. J.* **2010**, *4* (1), I1.
(122) Lombardo, A.; Roncaglioni, A.; Boriani, E.; Milan, C.; Benfenati, E. Assessment and Validation of the CAESAR Predictive Model for Bioconcentration Factor (BCF) in Fish. *Chem. Cent. J.* **2010**, *4* (1), S1.
(123) Mansouri, K.; Grulke, C. M.; Richard, A. M.; Judson, R. S.; Williams, A. J. An Automated Curation Procedure for Addressing Chemical Errors and Inconsistencies in Public Datasets Used in QSAR Modelling. *SAR QSAR Environ. Res.* **2016**, *27* (11), 911−937.
(124) Mansouri, K.; Grulke, C.; Richard, A.; Richard; Judson, S.; Williams, A. J.; Grulke, C.; Judson, R.; Kamel; Mansouri; Fourches, D.; Richard, A.; Grulke, C.; Williams, A. OPERA-Model for Octanol-Water Partition Coefficient. *Abstr. Pap. Am. Chem. Soc.* **2017**, *10* (10), 253.

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX

(125) Ferrari, T.; Cattaneo, D.; Gini, G.; Golbamaki Bakhtyari, N.; Manganaro, A.; Benfenati, E. Automatic Knowledge Extraction from Chemical Structures: The Case of Mutagenicity Prediction. *SAR QSAR Environ. Res.* **2013**, *24* (5), 365−383.
(126) Floris, M.; Manganaro, A.; Nicolotti, O.; Medda, R.; Mangiatordi, G. F.; Benfenati, E. A Generalizable Definition of Chemical Similarity for Read-Across. *J. Chem. Inf.* **2014**, *6* (1), 39.
(127) Bajusz, D.; Rácz, A.; Héberger, K. Why Is Tanimoto Index an Appropriate Choice for Fingerprint-Based Similarity Calculations? *J. Chem. Inf.* **2015**, *7* (1), 20.
(128) Guo, W.; Liu, J.; Dong, F.; Song, M.; Li, Z.; Khan, M. K. H.; Patterson, T. A.; Hong, H. Review of Machine Learning and Deep Learning Models for Toxicity Prediction. *Exp. Biol. Med.* **2023**, *248* (21), 1952−1973.
(129) Zubrod, J. P.; Galic, N.; Vaugeois, M.; Dreier, D. A. Bio-QSARs 2.0: Unlocking a New Level of Predictive Power for Machine Learning-Based Ecotoxicity Predictions by Exploiting Chemical and Biological Information. *Environ. Int.* **2024**, *186*, 108607.
(130) Mayr, A.; Klambauer, G.; Unterthiner, T.; Hochreiter, S. DeepTox: Toxicity Prediction Using Deep Learning. *Front. Environ. Sci.* **2016**, *3* (80), 00080.
(131) Reif, D. M.; Martin, M. T.; Tan, S. W.; Houck, K. A.; Judson, R. S.; Richard, A. M.; Knudsen, T. B.; Dix, D. J.; Kavlock, R. J. Endocrine Profiling and Prioritization of Environmental Chemicals Using ToxCast Data. *Environ. Health Perspect.* **2010**, *118* (12), 1714−1720.
(132) Gini, G.; Franchi, A. M.; Manganaro, A.; Golbamaki, A.; Benfenati, E. ToxRead: A Tool to Assist in Read across and Its Use to Assess Mutagenicity of Chemicals. *SAR QSAR Environ. Res.* **2014**, *25* (12), 999−1011.
(133) Viganò, E. L.; Colombo, E.; Raitano, G.; Manganaro, A.; Sommovigo, A.; Dorne, J. L. C.; Benfenati, E. Virtual Extensive Read-Across: A New Open-Access Software for Chemical Read-Across and Its Application to the Carcinogenicity Assessment of Botanicals. *Molecules* **2022**, *27* (19), 6605.
(134) Yu, Z.; Wu, Z.; Zhou, M.; Cao, K.; Li, W.; Liu, G.; Tang, Y. EDC-Predictor: A Novel Strategy for Prediction of Endocrine-Disrupting Chemicals by Integrating Pharmacological and Toxicological Profiles. *Environ. Sci. Technol.* **2023**, *57* (46), 18013−18025.
(135) Dimitrov, S. D.; Diderich, R.; Sobanski, T.; Pavlov, T. S.; Chankov, G. V.; Chapkanov, A. S.; Karakolev, Y. H.; Temelkov, S. G.; Vasilev, R. A.; Gerova, K. D.; Kuseva, C. D.; Todorova, N. D.; Mehmed, A. M.; Rasenberg, M.; Mekenyan, O. G. QSAR Toolbox - Workflow and Major Functionalities. *SAR QSAR Environ. Res.* **2016**, *27* (3), 203−219.
(136) Sushko, I.; Novotarskyi, S.; Körner, R.; Pandey, A. K.; Rupp, M.; Teetz, W.; Brandmaier, S.; Abdelaziz, A.; Prokopenko, V. V.; Tanchuk, V. Y.; Todeschini, R.; Varnek, A.; Marcou, G.; Ertl, P.; Potemkin, V.; Grishina, M.; Gasteiger, J.; Schwab, C.; Baskin, I. I.; Palyulin, V. A.; Radchenko, E. V.; Welsh, W. J.; Kholodovych, V.; Chekmarev, D.; Cherkasov, A.; Aires-de-Sousa, J.; Zhang, Q.-Y.; Bender, A.; Nigsch, F.; Patiny, L.; Williams, A.; Tkachenko, V.; Tetko, I. V. Online Chemical Modeling Environment (OCHEM): Web Platform for Data Storage, Model Development and Publishing of Chemical Information. *J. Comput. Aided Mol. Des.* **2011**, *25* (6), 533−554.
(137) Sushko, I.; Salmina, E.; Potemkin, V. A.; Poda, G.; Tetko, I. V. ToxAlerts: A Web Server of Structural Alerts for Toxic Chemicals and Compounds with Potential Adverse Reactions. *J. Chem. Inf. Model.* **2012**, *52* (8), 2310−2316.
(138) Oprisiu, I.; Novotarskyi, S.; Tetko, I. V. Modeling of Non-Additive Mixture Properties Using the Online CHEmical Database and Modeling Environment (OCHEM). *J. Chem. Inf.* **2013**, *5* (1), 4.
(139) Tetko, I. V.; Novotarskyi, S.; Sushko, I.; Ivanov, V.; Petrenko, A. E.; Dieden, R.; Lebon, F.; Mathieu, B. Development of Dimethyl Sulfoxide Solubility Models Using 163 000 Molecules: Using a Domain Applicability Metric to Select More Reliable Predictions. *J. Chem. Inf. Model.* **2013**, *53* (8), 1990−2000.
(140) Peets, P.; Wang, W.-C.; MacLeod, M.; Breitholtz, M.; Martin, J. W.; Kruve, A. MS2Tox Machine Learning Tool for Predicting the Ecotoxicity of Unidentified Chemicals in Water by Nontarget LC-HRMS. *Environ. Sci. Technol.* **2022**, *56* (22), 15508−15517.
(141) Arturi, K.; Hollender, J. Machine Learning-Based Hazard-Driven Prioritization of Features in Nontarget Screening of Environmental High-Resolution Mass Spectrometry Data. *Environ. Sci. Technol.* **2023**, *57* (46), 18067−18079.
(142) Vandenberg, L. N.; Colborn, T.; Hayes, T. B.; Heindel, J. J.; Jacobs, D. R., Jr; Lee, D.-H.; Shioda, T.; Soto, A. M.; vom Saal, F. S.; Welshons, W. V.; Zoeller, R. T.; Myers, J. P. Hormones and Endocrine-Disrupting Chemicals: Low-Dose Effects and Nonmonotonic Dose Responses. *Endocr. Rev.* **2012**, *33* (3), 378−455.
(143) Zhu, J.-J.; Yang, M.; Ren, Z. J. Machine Learning in Environmental Research: Common Pitfalls and Best Practices. *Environ. Sci. Technol.* **2023**, *57* (46), 17671−17689.
(144) Jia, X.; Wang, T.; Zhu, H. Advancing Computational Toxicology by Interpretable Machine Learning. *Environ. Sci. Technol.* **2023**, *57* (46), 17690−17706.
(145) OECD *Implementing the OECD AI Principles: Challenges and Best Practices*, 2022.
(146) Idakwo, G.; Luttrell, J.; Chen, M.; Hong, H.; Zhou, Z.; Gong, P.; Zhang, C. A Review on Machine Learning Methods for in Silico Toxicity Prediction. *J. Environ. Sci. Health Part C Environ. Health Sci.* **2018**, *36* (4), 169−191.
(147) Worth, A.; Fuart-Gatnik, M.; Lapenna, S.; Serafimova, R. Applicability of QSAR Analysis in the Evaluation of Developmental and Neurotoxicity Effects for the Assessment of the Toxicological Relevance of Metabolites and Degradates of Pesticide Active Substances for Dietary Risk Assessment. *EFSA Supporting Publ.* **2011**, *8* (6), 169E.
(148) More, S. J.; Bampidis, V.; Benford, D.; Bragard, C.; Halldorsson, T. I.; Hernández-Jerez, A. F.; Hougaard Bennekou, S.; Koutsoumanis, K. P.; Machera, K.; Naegeli, H.; Nielsen, S. S.; Schlatter, J. R.; Schrenk, D.; Silano, V.; Turck, D.; Younes, M.; Gundert-Remy, U.; Kass, G. E. N.; Kleiner, J.; Rossi, A. M.; Serafimova, R.; EFSA Scientific Committee; et al. Guidance on the Use of the Threshold of Toxicological Concern Approach in Food Safety Assessment. *EFSA J.* **2019**, *17* (6), No. e05708.
(149) Bloch, D.; Diel, P.; Epe, B.; Hellwig, M.; Lampen, A.; Mally, A.; Marko, D.; Villar Fernández, M. A.; Guth, S.; Roth, A.; Marchan, R.; Ghallab, A.; Cadenas, C.; Nell, P.; Vartak, N.; van Thriel, C.; Luch, A.; Schmeisser, S.; Herzler, M.; Landsiedel, R.; Leist, M.; Marx-Stoelting, P.; Tralau, T.; Hengstler, J. G. Basic Concepts of Mixture Toxicity and Relevance for Risk Evaluation and Regulation. *Arch. Toxicol.* **2023**, *97* (11), 3005−3017.
(150) Sigmund, G.; Arp, H. P. H.; Aumeier, B. M.; Bucheli, T. D.; Chefetz, B.; Chen, W.; Droge, S. T. J.; Endo, S.; Escher, B. I.; Hale, S. E.; Hofmann, T.; Pignatello, J.; Reemtsma, T.; Schmidt, T. C.; Schönsee, C. D.; Scheringer, M. Sorption and Mobility of Charged Organic Compounds: How to Confront and Overcome Limitations in Their Assessment. *Environ. Sci. Technol.* **2022**, *56* (8), 4702−4710.
(151) Henneberger, L.; Goss, K.-U. Environmental Sorption Behavior of Ionic and Ionizable Organic Chemicals. In *Reviews of Environmental Contamination and Toxicology*; de Voogt, P., Ed.; Springer International Publishing: Cham, 2019; Vol. *253*, pp 43−64..
(152) *European Parliament and Council of the European Union. Directive (EU) 2024/3019 of the European Parliament and of the Council of 27 November 2024 Concerning Urban Wastewater Treatment (Recast) (Text with EEA Relevance)*. Official Journal of the European Union 2024, No. L 202, 1−65.

https://doi.org/10.1021/acs.est.5c06790
Environ. Sci. Technol. XXXX, XXX, XXX−XXX