Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Toxicology

Supporting systems science through *in silico* applications: A focus on informing metabolic mechanisms

Jeremy A. Leonard

Abstract

Despite the advent of new *in silico* methodologies to characterize exposure and pharmacokinetic properties of environmental chemicals, numerous limitations remain in the understanding and predicting of metabolic processes. Metabolomics has arisen as means to better inform underlying mechanisms of metabolism and biological effects following chemical exposure. Metabolomics is applied for numerous purposes, including clinical diagnoses of disease, biomarker discovery, toxicological studies, and environmental monitoring. One of the primary challenges involved in metabolomics is identification of detectable metabolites, due to lack of chemical information in databases or reference spectra. Several computational tools, such as metabolite-predicting software and *in silico* spectral libraries, have been developed to address such issues. Herein, application of such tools is discussed to support environmental monitoring and nontargeted chemical analyses or approaches that are based on pathway frameworks designed to better inform chemical effects within living systems. As these tools advance over time and are able to provide more accurate prediction of metabolites, the list of identifiable chemicals within a given exposure medium will continue to grow. This, in turn, will improve confidence in exposure-based and hazard-based decision-making practices.

Addresses

Oak Ridge Institute for Science and Education, Oak Ridge, TN, 37381, United States

Corresponding author: Leonard, Jeremy A. (jaleonar@ncsu.edu)

Current Opinion in Toxicology 2019, **16:**1–8

This review comes from a themed issue on **Systems toxicology (2019)**

Available online 21 March 2019

For a complete overview see the Issue and the Editorial

https://doi.org/10.1016/j.cotox.2019.03.007

Keywords

Metabolomics, Environmental monitoring, Nontargeted analyses, Adverse outcome pathway, Aggregate exposure pathway, Pharmacokinetics.

In silico predictions of pharmacokinetic behaviors to support risk assessment

Ensuring chemical safety often requires information regarding the relationship between external concentrations and biologically effective internal concentrations [1]. This relationship is influenced by pharmacokinetic (PK) behaviors such as absorption, distribution, metabolism, and elimination (ADME). Because a large number of chemicals in production or currently on the market lack comprehensive exposure and PK information, a significant effort has been taken to predict these chemical-specific characteristics through *in silico* means. Many computational exposure tools require inputs of ingredient concentrations, exposure duration, consumer activity, and product use to generate their predictions [2–5]. Ingredient information for products can be found in repositories such as the United States Environmental Protection Agency's (US EPA) Chemical and Product Database [6,7]. Information regarding exposure routes and duration to certain population groups can be found in the US EPA's Exposure Factor Handbook [8], which is supported by an online tool that provides advice on exposure assessment methodologies (https://www.epa.gov/expobox). In addition, several databases are available that provide information about consumer activity and product use. Some examples are the European Food and Safety Authority's Comprehensive European Food Consumption Database (http://www.efsa.europa.eu/en/food-consumption/comprehensive-database) [9] and the US EPA's Consolidated Human Activity Database (https://www.epa.gov/healthresearch/consolidated-human-activity-database-chad-use-human-exposure-and-health-studies-and) that include product information and demographic metrics such as age. Despite this wealth of available information, exposure predictions can sometimes span an order of several magnitudes for some chemicals [10] as human behaviors vary widely and influence exposure variability more so than chemical information. In seeking to avoid unreliable exposure estimates, some studies have used physicochemical properties to predict exposure routes and the most susceptible exposure groups from among the general population [11,12]. Physicochemical properties are also often the basis behind predictions of absorption and distribution. For example, the number of hydrogen bond donors/acceptors and the octanol–water partition coefficient of a chemical can be used to estimate its oral bioavailability [13], and properties such as molecular weight and water solubility can be used to predict a chemical's ease of absorbance across the skin

[14]. Distribution across biological barriers such as the brain or placenta can also be predicted by input of physicochemical descriptors like molecular weight, polar surface area, and the octanol–water partition coefficient into quantitative structure activity relationship (QSAR) models [15–19].

Use of predictions to aid in characterization of clearance

Absorption and distribution can influence the ability of an active chemical to reach its intended biological target once in the body. Clearance can also impact the ability of an inactive parent chemical to reach its target in an active form, by mediating disappearance of that parent through transformation or excretion. Clearance can occur through biliary clearance and fecal excretion, via transport-mediated processes occurring along the apical and basolateral membranes of hepatocytes. Because bile is not easily collected in healthy individuals, estimation of biliary clearance is difficult despite several *in vivo* and *in vitro* methodologies, such as fecal recovery and sandwich-configuration hepatocytes [20,21]. Predictions of biliary clearance for poorly metabolized compounds are based on physicochemical properties such as molecular weight and polarizability [22]. Renal clearance also remains a major route of elimination, via both active transport and passive reabsorption mechanisms. Renal clearance can be predicted through knowledge of unbound chemical fraction in the plasma, dipole moments and electrical charge, molecule size, and physicochemical properties such as lipophilicity [23]. Studies conducted to predict these types of clearance mechanisms have resulted in the development of the Extended Clearance Classification System for the purposes of drug discovery, which is based on passive membrane permeability and physicochemical properties [24,25]. Advanced knowledge of a chemical's mechanism of clearance (i.e., hepatic, renal, or biliary) allows for better predicting its clearance rate via that mechanism, as the training set in the model can be constrained to chemicals only undergoing that specific mechanism [26]. Other routes of elimination can include exhalation, perspiration, and loss through breast milk, with these pathways also heavily dependent on chemical-specific properties such as volatility or lipophilicity.

It should be noted that a majority of clearance models have been developed using primarily pharmaceuticals within the training set, because of a greater availability of data for such compounds; applying these models to estimate clearance mechanisms or rates for environmental chemicals may result in domain of applicability issues. Often, clearance of environmental chemicals occurs primarily through phase I and phase II metabolism or a combination of renal and metabolic processes [10], and so determination of total clearance may provide the best estimate for such chemicals. In its simplest form, total clearance can be derived through knowledge of a chemical's volume of distribution (V_d), half life ($t_{1/2}$), and body weight (BW) according to the following equation:

$$CL_{total} = \frac{\ln(2)V_d}{t_{\frac{1}{2}}}\mathrm{BW}$$

Volume of distribution can be predicted using properties such as a chemical's lipophilicity, ionization state, and tissue–plasma partition coefficient at a given pH [27,28], and half-life has been extrapolated in humans based on observed half-life in rodents and a linear equation consisting of two constants and the chemical's octanol–water partition coefficient [29].

Opportunities and challenges associated with metabolomics approaches

The study of metabolites in cells, tissues, organs, or biofluids, especially those generated in response to a xenobiotic stressor, is known as metabolomics [30,31]. Metabolomics has arisen as a critical application of systems biology to better characterize the mechanisms underlying metabolism. One significant advantage of metabolomics is that a multivariate profile of metabolites can be generated upon biological perturbation, thus aiding in biomarker discovery. Such applications have been used in aquatic toxicology [32,33], drug discovery [34], clinical diagnoses of disease [35], and identification of metabolites. Analyses can involve use of nuclear magnetic resonance spectrometry, because of its nondestructive nature and high reproducibility of data [36]. More recently, in an effort to apply approaches with a higher sensitivity to resolve a greater number of compounds, gas chromatography or liquid chromatography, coupled with mass spectrometry, has risen in use [37]. Regardless of the method, it is combined with multivariate statistical approaches to better understand the complexity involved in the analysis of hundreds to thousands of low-molecular-weight metabolites [38]. Despite the numerous powerful technological tools and applications that are available to support metabolomics approaches, identification of metabolites remains a significant challenge. This limitation is often due to sparse or limited data available for a compound's structure, metabolic pathway, or reference spectra.

Using *in silico* tools to inform metabolic mechanisms in support of environmental monitoring and nontargeted chemical analyses

One recent application of metabolomics has been the identification of metabolites in biological and environmental media [39–41] to aid targeted and nontargeted environmental monitoring strategies. Traditional

monitoring of an environmental or biological medium involves extensive analysis of tens to hundreds of select compounds [42,43]. These compounds are often chosen primarily because of *a priori* knowledge of their presence in that particular medium and of adverse biological responses they might elicit in humans or other organisms upon exposure. As technical instrumentation has advanced, the lowering of detection limits and increase in measurement capabilities has enabled identification of a greater number of compounds in environmental and biological media over the years. This, in turn, has led to emerging data that can be used to support conventional and novel exposure tools and models, thus allowing for risk assessment to transition from a primarily hazard-based approach to one that equally considers chemical exposure prior to health response [44]. As a result, there has arisen an impetus to investigate exposure for the hundreds to thousands of chemicals in existence in order for exposure science to keep pace with 21st century toxicity testing. To achieve this goal, one strategy has involved development of nontargeted monitoring of animal tissue or biofluids to complement the more traditional targeting of select compounds [45]. This approach can allow for identification of heretofore-unknown compounds within a given environmental medium and lead to hypothesis generation in regards to what their presence might mean to organisms inhabiting that medium. It should be noted that the metabolites extracted from an animal's tissue can be derived through its own metabolic mechanisms or can be taken in as preformed compounds that have been generated through prior microbial or physicochemical processes (e.g., photolysis or hydrolysis). As metabolomics is a study of chemical and biological processes at a system level, its application to environmental monitoring of xenobiotic stress response is obvious. Metabolomics approaches are relevant also to identification of metabolites produced through microbial mechanisms occurring outside of an organism, because those metabolites are generated through biological processes. Although metabolomics is a means to investigate biological systems, the primary outcome of such an approach is characterization of metabolic mechanisms, and as such, the tools involved in this discipline can also be utilized to aid in the identification of metabolites present in abiotic media such as air, water, and dust [46–48].

Nontargeted analysis of abiotic media involves attempting to match spectra or formula determined experimentally against known compound formulas and masses found in online chemical databases such as ChemSpider (http://www.chemspider.com/Default.aspx), PubChem (https://pubchem.ncbi.nlm.nih.gov/) or the United States Environmental Protection Agency's Computational Toxicology Dashboard (CompTox; https://comptox.epa.gov/dashboard), as well as vendor or in-house databases that may or may not be available online. However, unmatched spectra or formulas can often result in a lack of detection of many chemicals that might actually be present in a medium, because of the absence of their empirical information in these databases [49]. One possibility is that the parent chemicals themselves may be absent from the searched chemical database. More often, it is the case that, despite the dual biological and chemical nature of metabolites, their information is also absent from these databases.

Given the large number of chemicals present throughout the environment and the much greater number of multiple transformation products for each chemical, it is no surprise that structures or masses are often not available for these 'dark matter' of the metabolome [50]. Fortunately, there are numerous computational tools that have been developed in the effort to facilitate identification of metabolites. Many of these platforms predict only sites of metabolism, whereas others extend their capabilities to also predict metabolites generated through any one of the many known metabolic pathways, especially from cytochrome P450–mediated processes. A few examples of such platforms include Meteor Nexus (Lhasa Limited), QSAR Toolbox (Organization for Economic Cooperation and Development), MetabolExpert (CompuDrug), MetaPrint2D-React [51], and ADMET Predictor (Simulations Plus, Inc.) One common downside with such *in silico* tools, however, is that all possible known reaction schemes are inputted into a library and so will result in prediction of any hypothetical metabolite known to be generated from that particular scheme. Another limitation is that library reaction schemes for a platform may be based on nonhuman metabolic mechanisms (e.g., rodent or microbial metabolism) [52]. Many of these *in silico* tools contain proprietary data, and recently one open-source platform was developed to predict a reasonable number of metabolites for single chemicals or chemical batches through knowledge of six reaction systems including both human and soil/aquatic microbial metabolism, with the reaction libraries obtained from empirical data within literature [53].

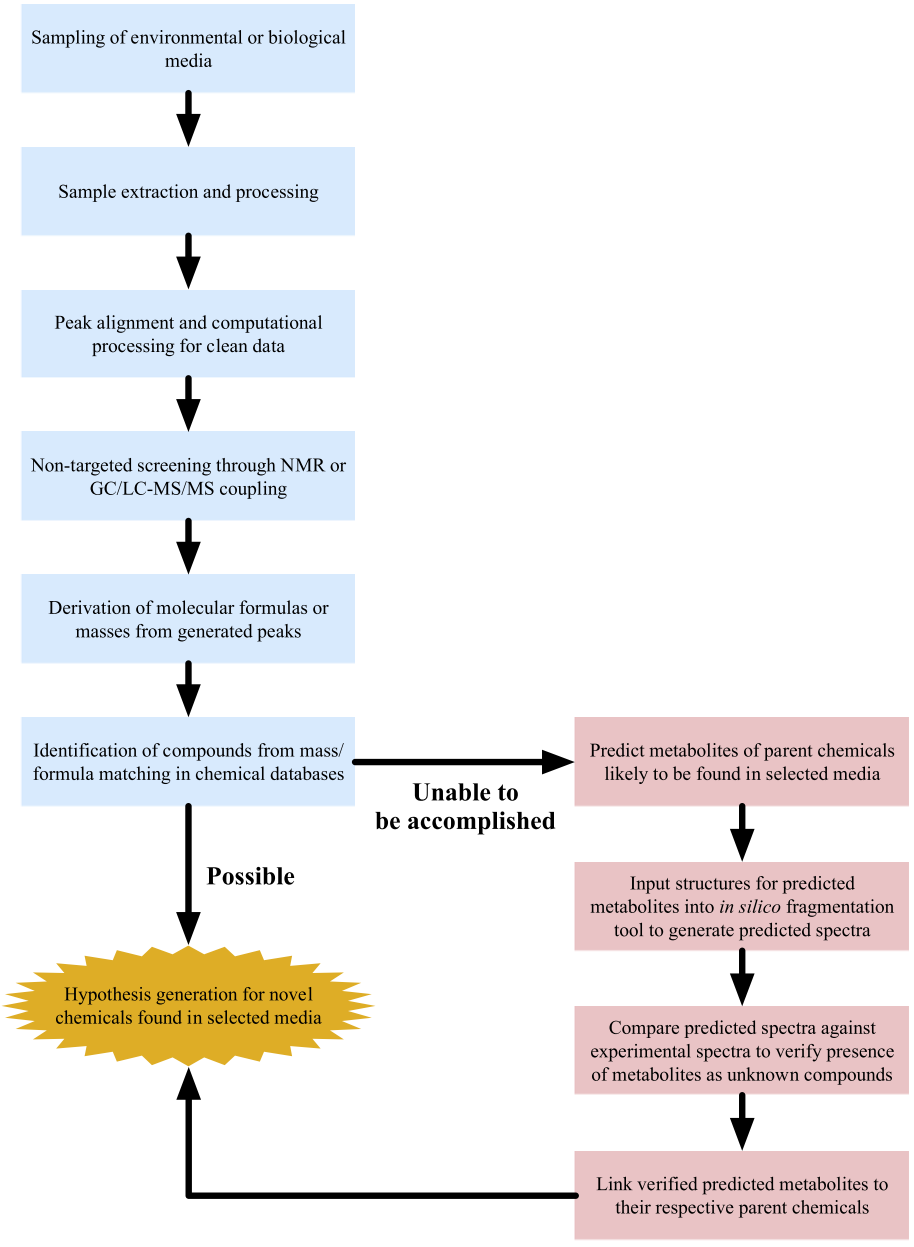
The presence of metabolites predicted by any of these platforms can be verified by comparing experimental spectra generated from extracts of the select medium to reference spectra contained in online databases. Some examples of these spectral libraries include METLIN [54], the Mass Spectral Library developed by the National Institute for Standards and Technology (http://nistmassspeclibrary.com/), and the Golm Metabolome Database [55]. Unfortunately, it is often the case that spectra for these metabolites will not be available, just as their general chemical information is not available within chemical databases. As a result, several *in silico* tools have been developed that contain predicted spectra based on how a chemical is expected to fragment in a mass spectrometer [56]. Such spectral predictive

tools include LipidBlast [57], MetFrag [58], MS-Finder [59], and the Competitive Fragmentation Modeling for Metabolite Identification tool [60]. The combined use of *in silico* applications that predict metabolites and generate spectra for those putative metabolites can complement nontargeted analysis (Figure 1). By combining both biological screening and nontargeted chemical analysis that apply similar technological tools, chemicals within multiple media types can be more effectively monitored and managed [61].

Expanding the utility of systems-based pathways in risk assessment through *in silico* characterization of metabolism

The Mode of Action framework was developed to aid in human risk assessment by establishing the relationship between a biological effect and exposure to a specific chemical or chemicals [62,63]. This was later followed by the chemical-agnostic Adverse Outcome Pathway (AOP) framework, which was developed primarily for ecosystem risk assessment and then expanded to human

Figure 1



Workflow of *in silico* approaches (red boxes) to complement nontarget screening (blue boxes) for identification of detectable compounds within biological and environmental media. NMR, nuclear magnetic resonance; GC, gas chromatography; LC, liquid chromatography; MS, mass spectrometry.

risk assessment applications [64–66]. Both frameworks operate by organizing chemical knowledge and data across a series of biologically connected events within a living system to allow for informing mechanism-based outcomes relevant to human and ecological health. The practicality of the AOP framework lies in its underlying premise that induction of the first key event within the series (i.e., the molecular initiating event) will eventually lead to a final adverse biological outcome, regardless of the chemical involved. More recently, this premise has led to the suggestion that the AOP framework can be integrated with a newly proposed Aggregate Exposure Pathway (AEP) framework. The AEP framework is designed to follow a chemical or nonchemical stressor from its site of origination to its site of biological action [67,68]. It can be assumed that, given a sufficient concentration, the site at which a chemical acts is analogous to the site at which the molecular initiating event of an AOP will be triggered. Thus, integration of these two frameworks allows for better understanding the entire source-to-outcome continuum for a given chemical [69].

Often, the action of one chemical at a specific site is insufficient for triggering a molecular initiating event, and so an AOP might not be assumed possible for that chemical. However, multiple chemicals acting at the same site can lead to an adverse outcome, and so it is critical that all those potential chemicals are considered. Many times, these chemicals will include active metabolites of an inactive parent compound [51]. Predicting and verifying these metabolites through the aforementioned *in silico* approaches allows for identifying additional chemicals that can perturb a given biological target. Alternatively, these metabolites may act upon other molecular sites and lead to different adverse outcomes that might not otherwise be identified if not accounting for those metabolites.

Conclusions

Numerous technological tools have been developed to address the growing number of chemicals released to market on an annual basis [70] and to advance the field of exposure science [44]. A lack of available chemical-specific data for many chemicals has resulted in development of *in silico* methodologies to investigate exposure and PK characteristics for evaluating their safety risks. While a majority of PK behaviors can be predicted through knowledge of physicochemical properties, characterization of clearance, especially metabolic mechanisms, remains a significant challenge. Investigating the metabolome of a living organism exposed to environmental chemicals can provide insight into resulting phenotypic biological effects, and metabolomics tools, such as metabolite prediction applications, offer great potential for environmental monitoring. The integrated chromatography–mass spectrometry approaches and tools used for metabolomics studies are also easily transferrable to nontargeted analysis of abiotic media for investigators wishing to refine their chemical detection practices. Many major obstacles remaining for metabolomics to meet its true potential, such as lack of spectral reference libraries and inability to identify detectable metabolites, are being addressed through development of *in silico* tools such as metabolite-predicting software and predictive spectral libraries. Complementing nontargeted environmental analysis with bioanalytical screening will help to identify chemicals that previously may have been unable to be detected in a given medium, even if their presence is likely. The Toxic Substances Control Act involves regulation of novel or existing chemicals (https://www.epa.gov/tsca-inventory), and the ability to detect and identify chemicals that were previously unable to be identified in a particular medium can significantly aid in facilitating regulation of such chemicals. These *in silico* tools that help to advance exposure science and that help to fill data gaps in a chemical's AEP framework can also help to improve confidence in risk-based decisions applying knowledge of an AOP.

Disclaimer

The United States Environmental Protection Agency has provided administrative review and has approved this article for publication. The views expressed herein are those of the author and do not necessarily reflect the views of the agency.

Acknowledgements

Funding for Dr. Leonard was provided by the Oak Ridge Institute for Science and Education Research Participation Program at the US EPA. Special thanks to Drs. Seth Newton and Peter Egeghy for their critical review of this manuscript.

Conflicts of interest statement

Nothing declared.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Groh KJ, Carvalho RN, Chipman JK, Denslow ND, Halder M, Murphy CA, Roelofs D, Rolaki A, Schirmer K, Watanabe KH: **Development and application of the adverse outcome pathway framework for understanding and predicting chronic toxicity: I. Challenges and research needs in ecotoxicology**. *Chemosphere* 2015, **120**:764–777. https://doi.org/10.1016/j.chemosphere.2014.09.068.

2. Wambaugh JF, Wang A, Dionisio KL, Frame A, Egeghy P, Judson R, Setzer RW: **High throughput heuristics for prioritizing human exposure to environmental chemicals**. *Environ Sci Technol* 2014, **48**:12760–12767. https://doi.org/10.1021/es503583j.

3. European Food and Safety Authority: **Overview of existing methodologies for the estimation of non-dietary exposure to chemicals from the use of consumer products and via the environment**. *EFSA J* 2016, **14**:e04525. https://doi.org/10.2903/j.efsa.2016.4525.

4. Dudzina T, Delmaar CJE, Biesterbos JWH, Bakker MI, Bokkers BGH, Scheepers PTJ, van Engelen JGM, Hungerbuehler K, von Goetz N: **The probabilistic aggregate consumer exposure model (PACEM): validation and comparison to a lower-tier assessment for the cyclic siloxane D5**. *Environ Int* 2015, **79**:8–16. https://doi.org/10.1016/j.envint.2015.03.006.

5. Isaacs KK, Glen WG, Egeghy P, Goldsmith M-R, Smith L, Vallero D, Brooks R, Grulke CM, Özkaynak H: **SHEDS-HT: an integrated probabilistic exposure model for prioritizing exposures to chemicals with near-field and dietary sources**. *Environ Sci Technol* 2014, **48**:12750–12759. https://doi.org/10.1021/es502513w.

6. Goldsmith M-R, Grulke CM, Brooks RD, Transue TR, Tan YM, Frame A, Egeghy PP, Edwards R, Chang DT, Tornero-Velez R, Isaacs K, Wang A, Johnson J, Holm K, Reich M, Mitchell J, Vallero DA, Phillips L, Phillips M, Wambaugh JF, Judson RS, Buckley TJ, Dary CC: **Development of a consumer product ingredient database for chemical exposure screening and prioritization**. *Food Chem Toxicol Int J Publ Br Ind Biol Res Assoc* 2014, **65**:269–279. https://doi.org/10.1016/j.fct.2013.12.029.

7. Dionisio KL, Frame AM, Goldsmith M-R, Wambaugh JF, Liddell A, Cathey T, Smith D, Vail J, Ernstoff AS, Fantke P, Jolliet O, Judson RS: **Exploring consumer exposure pathways and patterns of use for chemicals in the environment**. *Toxicol Rep* 2015, **2**:228–237. https://doi.org/10.1016/j.toxrep.2014.12.009.

8. United States Environmental Protection Agency: *Exposure factors Handbook 2011 edition (final report)*. Washington, DC, USA: U.S. Environmental Protection Agency; 2011. https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NCEA&dirEntryId=236252. Accessed 7 March 2019.

9. European Food Safety Authority: **Use of the EFSA comprehensive European Food Consumption Database in exposure assessment**. *EFSA J* 2011, **9**. https://doi.org/10.2903/j.efsa.2011.2097.

10. Leonard JA, Sobel Leonard A, Chang DT, Edwards S, Lu J, Scholle S, Key P, Winter M, Isaacs K, Tan Y-M: **Evaluating the impact of uncertainties in clearance and exposure when prioritizing chemicals screened in high-throughput assays**. *Environ Sci Technol* 2016, **50**:5961–5971. https://doi.org/10.1021/acs.est.6b00374.

11. Dellarco M, Zaleski R, Gaborek BJ, Qian H, Bellin CA, Egeghy P, Heard N, Jolliet O, Lander DR, Sunger N, Stylianou KS, Tanir JY: **Using exposure bands for rapid decision making in the RISK21 tiered exposure assessment**. *Crit Rev Toxicol* 2017, **47**:317–341. https://doi.org/10.1080/10408444.2016.1270255.

12. Greggs W, Burns T, Egeghy P, Embry MR, Fantke P, Gaborek B, Heine L, Jolliet O, Lee C, Muir D, Plotzke K, Rinkevich J, Sunger N, Tanir JY, Whittaker M: **Qualitative approach to comparative exposure in alternatives assessment: comparative Exposure in Alternatives Assessment**. *Integr Environ Assess Manag* 2018. https://doi.org/10.1002/ieam.4070.

13. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ: **Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings**. *Adv Drug Deliv Rev* 1997, **23**:3–25. https://doi.org/10.1016/S0169-409X(96)00423-1.

14. Magnusson BM, Anissimov YG, Cross SE, Roberts MS: **Molecular size as the main determinant of solute maximum flux across the skin**. *J Invest Dermatol* 2004, **122**:993–999. https://doi.org/10.1111/j.0022-202X.2004.22413.x.

15. Toropov AA, Toropova AP, Beeg M, Gobbi M, Salmona M: **QSAR model for blood-brain barrier permeation**. *J Pharmacol Toxicol Methods* 2017, **88**:7–18. https://doi.org/10.1016/j.vascn.2017.04.014.

16. Takaku T, Nagahori H, Sogame Y, Takagi T: **Quantitative structure–activity relationship model for the fetal–maternal blood concentration ratio of chemicals in humans**. *Biol Pharm Bull* 2015, **38**:930–934. https://doi.org/10.1248/bpb.b14-00883.

17. Ma X, Chen C, Yang J: **Predictive model of blood-brain barrier penetration of organic compounds**. *Acta Pharmacol Sin* 2005, **26**:500–512. https://doi.org/10.1111/j.1745-7254.2005.00068.x.

18. Geldenhuys WJ, Mohammad AS, Adkins CE, Lockman PR: **Molecular determinants of blood–brain barrier permeation**. *Ther Deliv* 2015, **6**:961–971. https://doi.org/10.4155/tde.15.32.

19. Zhang L, Zhu H, Oprea TI, Golbraikh A, Tropsha A: **QSAR modeling of the blood–brain barrier permeability for diverse organic compounds**. *Pharm Res* 2008, **25**:1902. https://doi.org/10.1007/s11095-008-9609-0.

20. Ghibellini G, Leslie EM, Brouwer KLR: **Methods to evaluate biliary excretion of drugs in humans: an updated review**. *Mol Pharm* 2006, **3**:198–211. https://doi.org/10.1021/mp060011k.

21. Abe K, Bridges AS, Brouwer KLR: **Use of sandwich-cultured human hepatocytes to predict biliary clearance of angiotensin II receptor blockers and HMG-CoA reductase inhibitors**. *Drug Metab Dispos* 2009, **37**:447–452. https://doi.org/10.1124/dmd.108.023465.

22. Hosey CM, Broccatelli F, Benet LZ: **Predicting when biliary excretion of parent drug is a major route of elimination in humans**. *AAPS J* 2014, **16**:1085–1096. https://doi.org/10.1208/s12248-014-9636-1.

23. Dave RA, Morris ME: **Quantitative structure-pharmacokinetic relationships for the prediction of renal clearance in humans**. *Drug Metab Dispos* 2015, **43**:73–81. https://doi.org/10.1124/dmd.114.059857.

24. El-Kattan AF, Varma MV, Steyn SJ, Scott DO, Maurer TS, Bergman A: **Projecting ADME behavior and drug-drug interactions in early discovery and development: application of the extended clearance classification system**. *Pharm Res* 2016, **33**:3021–3030. https://doi.org/10.1007/s11095-016-2024-z.

25. * Varma MV, Steyn SJ, Allerton C, El-Kattan AF: **Predicting clearance mechanism in drug discovery: extended clearance classification system (ECCS)**. *Pharm Res* 2015, **32**:3785–3802. https://doi.org/10.1007/s11095-015-1749-4.

26. * Lombardo F, Obach RS, Varma MV, Stringer R, Berellini G: **Clearance mechanism assignment and total clearance prediction in human based upon in silico models**. *J Med Chem* 2014, **57**:4397–4405. https://doi.org/10.1021/jm500436v.

27. Freitas AA, Limbu K, Ghafourian T: **Predicting volume of distribution with decision tree-based regression methods using predicted tissue:plasma partition coefficients**. *J Cheminf* 2015, **7**:6. https://doi.org/10.1186/s13321-015-0054-x.

28. Lombardo F, Obach RS, Shalaeva MY, Gao F: **Prediction of volume of distribution values in humans for neutral and basic drugs using physicochemical measurements and plasma protein binding data**. *J Med Chem* 2002, **45**:2867–2876. https://doi.org/10.1021/jm0200409.

29. Sarver JG, White D, Erhardt P, Bachmann K: **Estimating xenobiotic half-lives in humans from rat data: influence of log P**. *Environ Health Perspect* 1997, **105**:1204–1209.

30. * Peng B, Li H, Peng X-X: **Functional metabolomics: from biomarker discovery to metabolome reprogramming**. *Protein Cell* 2015, **6**:628–637. https://doi.org/10.1007/s13238-015-0185-x.

31. ** Bundy JG, Davey MP, Viant MR: **Environmental metabolomics: a critical review and future perspectives**. *Metabolomics* 2008, **5**:3. https://doi.org/10.1007/s11306-008-0152-0.

32. ** Viant M: **Metabolomics of aquatic organisms: the new 'omics' on the block**. *Mar Ecol Prog Ser* 2007, **332**:301–306. https://doi.org/10.3354/meps332301.

33. Skelton DM, Ekman DR, Martinović-Weigelt D, Ankley GT, Villeneuve DL, Teng Q, Collette TW: **Metabolomics for in situ environmental monitoring of surface waters impacted by contaminants from both point and nonpoint sources**. *Environ Sci Technol* 2014, **48**:2395–2403. https://doi.org/10.1021/es404021f.

34. ** Wishart DS: **Applications of metabolomics in drug discovery and development**. *Drugs RD* 2008, **9**:307–322. https://doi.org/10.2165/00126839-200809050-00002.

35. Clish CB: **Metabolomics: an emerging but powerful tool for precision medicine**. *Cold Spring Harb Mol Case Stud* 2015, **1**. https://doi.org/10.1101/mcs.a000588.

36. Markley JL, Brüschweiler R, Edison AS, Eghbalnia HR, Powers R, Raftery D, Wishart DS: **The future of NMR-based metabolomics**. *Curr Opin Biotechnol* 2017, **43**:34–40. https://doi.org/10.1016/j.copbio.2016.08.001.

37. Dettmer K, Aronov PA, Hammock BD: **Mass spectrometry-based metabolomics**. *Mass Spectrom Rev* 2007, **26**:51–78. https://doi.org/10.1002/mas.20108.

38. Viant MR, Sommer U: **Mass spectrometry based environmental metabolomics: a primer and review**. *Metabolomics* 2013, **9**:144–158. https://doi.org/10.1007/s11306-012-0412-x.

39. Li S, Todor A, Luo R: **Blood transcriptomics and metabolomics for personalized medicine**. *Comput Struct Biotechnol J* 2015, **14**:1–7. https://doi.org/10.1016/j.csbj.2015.10.005.

40. Lin CY, Wu H, Tjeerdema RS, Viant MR: **Evaluation of metabolite extraction strategies from tissue samples using NMR metabolomics**. *Metabolomics* 2007, **3**:55–67. https://doi.org/10.1007/s11306-006-0043-1.

41. Swenson TL, Jenkins S, Bowen BP, Northen TR: **Untargeted soil metabolomics methods for analysis of extractable organic matter**. *Soil Biol Biochem* 2015, **80**:189–198. https://doi.org/10.1016/j.soilbio.2014.10.007.

42. Velázquez-Gómez M, Hurtado-Fernández E, Lacorte S: **Differential occurrence, profiles and uptake of dust contaminants in the Barcelona urban area**. *Sci Total Environ* 2019, **648**:1354–1370. https://doi.org/10.1016/j.scitotenv.2018.08.058.

43. Zhou M, Ford B, Lee D, Tindula G, Huen K, Tran V, Bradman A, Gunier R, Eskenazi B, Nomura DK, Holland N: **Metabolomic markers of phthalate exposure in plasma and urine of pregnant women**. *Front Public Health* 2018, **6**:298. https://doi.org/10.3389/fpubh.2018.00298.

44. ** National Academies of Sciences, Engineering, and Medicine; Division on Earth and Life Studies; Board on Environmental Studies and Toxicology; Committee on Incorporating 21st Century Science into Risk-Based Evaluations. **Advances in exposure science**. In *Using 21st Century Science to Improve Risk-Related Evaluations*. Washington, D.C.: National Academies Press; 2017. https://www.ncbi.nlm.nih.gov/books/NBK424978/. Accessed 11 September 2017.

45. Heffernan AL, Gómez-Ramos MM, Gaus C, Vijayasarathy S, Bell I, Hof C, Mueller JF, Gómez-Ramos MJ: **Non-targeted, high resolution mass spectrometry strategy for simultaneous monitoring of xenobiotics and endogenous compounds in green sea turtles on the Great Barrier Reef**. *Sci Total Environ* 2017, **599–600**:1251–1262. https://doi.org/10.1016/j.scitotenv.2017.05.016.

46. Lee J, Hong S, Kwon B-O, Cha SA, Jeong H-D, Chang WK, Ryu J, Giesy JP, Khim JS: **Integrated assessment of persistent toxic substances in sediments from Masan Bay, South Korea: comparison between 1998 and 2014**. *Environ Pollut* 2018, **238**:317–325. https://doi.org/10.1016/j.envpol.2018.02.064. Barking Essex 1987.

47. Verkh Y, Rozman M, Petrovic M: **A non-targeted high-resolution mass spectrometry data analysis of dissolved organic matter in wastewater treatment**. *Chemosphere* 2018, **200**:397–404. https://doi.org/10.1016/j.chemosphere.2018.02.095.

48. Newton SR, McMahen RL, Sobus JR, Mansouri K, Williams AJ, McEachran AD, Strynar MJ: **Suspect screening and non-targeted analysis of drinking water using point-of-use filters**. *Environ Pollut* 2018, **234**:297–306. https://doi.org/10.1016/j.envpol.2017.11.033. Barking Essex 1987.

49. McEachran AD, Sobus JR, Williams AJ: **Identifying known unknowns using the US EPA's CompTox Chemistry Dashboard**. *Anal Bioanal Chem* 2017, **409**:1729–1735. https://doi.org/10.1007/s00216-016-0139-z.

50. * da Silva RR, Dorrestein PC, Quinn RA: **Illuminating the dark matter in metabolomics**. *Proc Natl Acad Sci U S A* 2015, **112**:12549–12550. https://doi.org/10.1073/pnas.1516878112.

51. Piechota P, Cronin MTD, Hewitt M, Madden JC: **Pragmatic approaches to using computational methods to predict xenobiotic metabolism**. *J Chem Inf Model* 2013, **53**:1282–1293. https://doi.org/10.1021/ci400050v.

52. Leonard JA, Stevens C, Mansouri K, Chang D, Pudukodu H, Smith S, Tan Y-M: **A workflow for identifying metabolically active chemicals to complement in vitro toxicity screening**. *Comput Toxicol* 2017. https://doi.org/10.1016/j.comtox.2017.10.003.

53. Feunang Djoumbou Y, Fiamoncini J, de la Fuente A, Manach C, Greiner R, Wishart D, BioTransformer: **A comprehensive computational tool for small molecule metabolism prediction and metabolite identification**. *J Cheminf* 2018. http://biotransformer.ca/.

54. Smith CA, Maille GO, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G: **METLIN: a metabolite mass spectral database**. *Ther Drug Monit* 2005, **27**:747–751. https://doi.org/10.1097/01.ftd.0000179845.53213.39.

55. Hummel J, Selbig J, Walther D, Kopka J: **The Golm Metabolome Database: a database for GC-MS based metabolite profiling**. In *Metabolomics powerful tool Syst. Biol.*. Edited by Nielsen J, Jewett MC, Berlin, Heidelberg: Springer Berlin Heidelberg; 2007:75–95. https://doi.org/10.1007/4735_2007_0229.

56. Tsugawa H: **Advances in computational metabolomics and databases deepen the understanding of metabolisms**. *Curr Opin Biotechnol* 2018, **54**:10–17. https://doi.org/10.1016/j.copbio.2018.01.008.

57. Kind T, Liu K-H, Lee DY, DeFelice B, Meissen JK, Fiehn O: **LipidBlast *in silico* tandem mass spectrometry database for lipid identification**. *Nat Methods* 2013, **10**:755–758. https://doi.org/10.1038/nmeth.2551.

58. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S: **MetFrag relaunched: incorporating strategies beyond in silico fragmentation**. *J Cheminf* 2016, **8**. https://doi.org/10.1186/s13321-016-0115-9.

59. Tsugawa H, Kind T, Nakabayashi R, Yukihira D, Tanaka W, Cajka T, Saito K, Fiehn O, Arita M: **Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software**. *Anal Chem* 2016, **88**:7946–7958. https://doi.org/10.1021/acs.analchem.6b00770.

60. Allen F, Pon A, Wilson M, Greiner R, Wishart D: **CFM-ID: a web server for annotation, spectrum prediction and metabolite identification from tandem mass spectra**. *Nucleic Acids Res* 2014, **42**:W94–W99. https://doi.org/10.1093/nar/gku436.

61. Maruya KA, Dodder NG, Mehinto AC, Denslow ND, Schlenk D, Snyder SA, Weisberg SB: **A tiered, integrated biological and chemical monitoring framework for contaminants of emerging concern in aquatic ecosystems**. *Integr Environ Assess Manag* 2016, **12**:540–547. https://doi.org/10.1002/ieam.1702.

62. * (Bette) Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, Longfellow DG, Pastoor T, Seed J, Patton DE: **A framework for human relevance analysis of information on carcinogenic modes of action**. *Crit Rev Toxicol* 2003, **33**:591–653. https://doi.org/10.1080/713608373.

63. (Bette) Meek ME, Palermo CM, Bachman AN, North CM, Jeffrey Lewis R: **Mode of action human relevance (species concordance) framework: evolution of the Bradford Hill considerations and comparative analysis of weight of evidence**. *J Appl Toxicol* 2014, **34**:595–606. https://doi.org/10.1002/jat.2984.

64. ** Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrrano JA, Tietge JE, Villeneuve DL: **Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment**. *Environ Toxicol Chem SETAC* 2010, **29**:730–741. https://doi.org/10.1002/etc.34.

65. Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, Landesmann B, Lettieri T, Munn S, Nepelska M, Ottinger MA, Vergauwen L, Whelan M: **Adverse outcome pathway (AOP) development I: strategies and principles**. *Toxicol Sci Off J Soc Toxicol* 2014, **142**:312–320. https://doi.org/10.1093/toxsci/kfu199.

66. Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, Landesmann B, Lettieri T, Munn S, Nepelska M, Ottinger MA, Vergauwen L, Whelan M: **Adverse outcome pathway development II: best practices**. *Toxicol Sci* 2014, **142**:321–330. https://doi.org/10.1093/toxsci/kfu200.

67. ** Teeguarden JG, Tan Y-M, Edwards SW, Leonard JA, Anderson KA, Corley RA, Kile ML, Simonich SM, Stone D, Tanguay RL, Waters KM, Harper SL, Williams DE: **Completing the link between exposure science and toxicology for improved environmental health decision making: the aggregate exposure pathway framework**. *Environ Sci Technol* 2016, **50**:4579–4586. https://doi.org/10.1021/acs.est.5b05311.

68. * Tan Y-M, Leonard JA, Edwards S, Teeguarden J, Egeghy P: **Refining the aggregate exposure pathway**. *Environ Sci Process Impact* 2018, **20**:428–436. https://doi.org/10.1039/C8EM00018B.

69. Tan Y-M, Leonard JA, Edwards S, Teeguarden J, Paini A, Egeghy P: **Aggregate exposure pathways in support of risk assessment**. *Curr Opin Toxicol* 2018, **9**:8–13. https://doi.org/10.1016/j.cotox.2018.03.006.

70. Egeghy PP, Judson R, Gangwal S, Mosher S, Smith D, Vail J, Cohen Hubal EA: **The exposure data landscape for manufactured chemicals**. *Sci Total Environ* 2012, **414**:159–166. https://doi.org/10.1016/j.scitotenv.2011.10.046.