LJMU Research Online

Zwickl, C and Cronin, MTD

Principles and Procedures for Assessment of Acute Toxicity Incorporating In Silico Methods

https://researchonline.ljmu.ac.uk/id/eprint/17239/

Article

For more information please contact researchonline@ljmu.ac.uk

Principles and Procedures for Assessment of Acute Toxicity Incorporating *In Silico* Methods

Craig M. Zwickl[a], Jessica C. Graham[b], Robert A. Jolly[c], Arianna Bassan[d], Ernst Ahlberg[e,f], Alexander Amberg[g], Lennart T. Anger[b], Lisa Beilke[h], Phillip Bellion[i], Alessandro Brigo[j], Heather Burleigh-Flayer[k], Mark T.D. Cronin[l], Amy A. Devlin[m], Trevor Fish[n], Susanne Glowienke[o], Kamila Gromek[p], Agnes Karmaus[q], Ray Kemper[r], Sunil Kulkarni[s], Elena Lo Piparo[t], Federica Madia[u], Matthew Martin[v], Melisa Masuda Herrera[w], Britt L. McAtee[k], Jordi Mestres[x,y], Lawrence Milchak[z], Chandrika Moudgal[b], Moiz Mumtaz[aa], Wolfgang Muster[j], Louise Neilson[bb], Grace Patlewicz[cc], Alexandre Paulino[dd], Alessandra Roncaglioni[ee], Patricia Ruiz[ff], David T. Szabo[k], Jean-Pierre Valentin[gg], Ioanna Vardakou[hh], David Woolley[ii], Glenn J. Myatt[jj]*

a)   Transendix LLC, Indianapolis, IN 46229, USA
b)   Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA
c)   Eli Lilly and Company, Indianapolis, IN 46285, USA
d)   Innovatune, Via Giulio Zanon 130/D, 35129 Padova, Italy
e)   Universal Prediction AB, Gothenburg, Sweden
f)   Department of Pharmaceutical Biosciences, Uppsala University, Uppsala, Sweden
g)   Sanofi, R&D Preclinical Safety Frankfurt, Industriepark Hoechst, D-65926 Frankfurt am Main, Germany
h)   Toxicology Solutions, Inc., 10531 4S Commons Dr. #594, San Diego, CA 92127, USA
i)   Boehringer Ingelheim Vetmedica GmbH, Binger Str. 173, 55216 Ingelheim am Rhein, Germany
j)   Roche Pharmaceutical Research & Early Development, Roche Innovation Center Basel, Grenzacherstrasse 124, 4070, Basel, Switzerland
k)   PPG Industries, Monroeville, PA 15146, USA
l)   Liverpool John Moores University, Liverpool, L3 3AF, UK
m)   FDA Center for Drug Evaluation and Research, Silver Spring, MD 20993, USA
n)   Nelson Laboratories, Salt Lake City, Utah, USA
o)   Novartis, Basel, Switzerland
p)   GlaxoSmithKline, Munich, Germany
q)   Integrated Laboratory Systems, LLC, Morrisville, NC, USA
r)   Nuvalent, Cambridge, MA 02142, USA
s)   Healthy Environments and Consumer Safety Branch, Health Canada / Government of Canada
t)   Chemical Food Safety Group, Nestlé Research, Lausanne, Switzerland
u)   European Commission, Joint Research Centre (JRC), Ispra, Italy
v)   Pfizer Worldwide Research & Development, Groton, CT USA
w)   Gilead Sciences, 333 Lakeside Drive, Foster City, CA 94404
x)   IMIM Institut Hospital del Mar d'Investigacions Mèdiques and Universitat Pompeu Fabra, Doctor Aiguader 88, Parc de Recerca Biomèdica, 08003 Barcelona, Spain
y)   Chemotargets SL, Baldiri Reixac 4, Parc Científic de Barcelona, 08028 Barcelona, Spain
z)   3M Company, St. Paul, MN
aa)   Office of the Associate Director for Science (OADS), Agency for Toxic Substances and Disease Registry (ATSDR), Centers for Disease Control and Prevention (CDC), Atlanta, GA 30333, USA
bb)   JT International SA, Geneva, Switzerland

cc) Centre for Computational Toxicology and Exposure (CCTE), US Environmental Protection Agency, 46 Research Triangle Park, NC, USA

Corresponding author:

E-mail address: glenn.myatt@instem.com (G.J. Myatt).

Abstract

Acute toxicity in silico models are being used to support an increasing number of application areas including (1) product research and development, (2) product approval and registration as well as (3) the transport, storage and handling of chemicals. The adoption of such models is being hindered, in part, because of a lack of guidance describing how to perform and document an in silico analysis. To address this issue, a framework for in acute toxicity hazard assessment is proposed. This framework combines results from different sources including in silico methods and in vitro or in vivo experiments. In silico methods that can assist the prediction of in vivo outcomes (i.e., LD50) are analyzed concluding that predictions obtained using in silico approaches are now well-suited for reliably supporting assessment of LD50-based acute toxicity for the purpose of the Globally Harmonized System (GHS) classification. A general overview is provided of the endpoints from in vitro studies commonly evaluated for predicting acute toxicity (e.g., cytotoxicity/cytolethality as well as assays targeting specific mechanisms). The increased understanding of pathways and key triggering mechanisms underlying toxicity and the increased availability of in vitro data allow for a shift away from assessments solely based on endpoints such as LD50, to mechanism-based endpoints that can be accurately assessed in vitro or by using in silico prediction models. This paper also highlights the importance of an expert review of all available information using weight-of-evidence considerations and illustrates, using a series of diverse practical use cases, how in silico approaches support the assessment of acute toxicity.

1. Introduction

Acute toxicity encompasses any one or more of that set of adverse events occurring after a single or several closely-spaced exposure(s) to an agent (e.g., chemical), and which usually occurs within 24 h, but that can take up to several weeks to manifest [1]. Toxicity can be localized whereby the adverse effect is limited to the area immediately proximal to a relatively confined site of initial exposure, or systemic, where the effect is widespread and can involve tissues and organs at locations distant from the site of initial exposure. Although systemic toxicity has the potential to manifest in any number of anatomical locations throughout the body, major effect(s) can be confined to a particular organ system or tissue that is uniquely sensitive to the toxicologic agent (i.e., the target organ of toxicity) [2].

The ubiquitous use of chemical substances in manufacturing, fabrication, and/or processing of all types and amounts of natural and synthetic materials including therapeutics, consumer products and goods, agricultural and veterinary products, food additives, devices, building materials, etc., together with the accompanying need to handle, transport, store, recycle, and dispose of chemicals have all contributed to what are essentially unavoidable opportunities for people to be exposed to chemicals at some level. This reality spotlights the need for increased scientific understanding and promotion of public awareness regarding possible adverse health effects from exposure to a chemical agent, including acute toxicity. As a result, hazard identification and assessment activities have relevance in many types of settings (e.g., product registration, occupational health and safety, environmental protection, and public health and safety) and reach across most industries. In recent years, classification schemes have been developed to standardize and communicate hazard information derived from these assessments [3].

First introduced in the 1920 s by Trevan [4], several standardized designs of the in vivo acute lethality test, which measures the median lethal dose (LD50) in animals (generally rodents), have been used as a 'gold standard' for evaluating the acute toxicity of chemicals. The basic design involves quantitatively estimating the dose (e.g., mg of substance administered per kg body weight) resulting in morbidity or lethality to 50 percent of the test animals over a two-week period following administration of either a single dose or multiple doses within a 24-hour period [5]. During the last several decades, design modifications have been introduced that aimed to minimize the number of animals used for testing without sacrificing the scientific integrity of the assay [6].

Decades of experience with the LD50 has demonstrated its usefulness for several purposes, including chemical screening and triaging compounds for further study, identifying starting doses for longer-term in vivo studies,

establishing exposure limits, comparing the toxicologic liability across chemicals, and identifying and classifying degrees of hazard.

Despite the usefulness of LD50 rodent testing, ethical considerations primarily centered on dosing animals to the point of mortality and/or morbidity, have provided a strong motivation to identify and validate alternative (modified design, non-mammalian, in vitro, in silico) testing approaches. As a result, reliance on studies in zebrafish, nematodes, and fruit flies, in vitro assay systems based on general readouts such as cytolethality or those based on mechanistic understanding, and in silico approaches, which may include applied informatics and/or the use of computational models, continue to gain traction across multiple sectors [1,7].

Generally, computational in silico approaches include a combination of techniques that may rely on expert knowledge, rule-based approaches, statistical-based methods, and/or read-across strategies. Increased availability of high-quality experiment-derived data for a variety of endpoints, can be leveraged to train in silico prediction models and continues to drive new opportunities for applying computational approaches more widely to the evaluation of chemico-biological interactions, including toxicity.

Although the use of in silico models is on the rise in many sectors, their wider acceptance may be hampered by a general lack of sufficiently detailed guidance, for example with respect to justifying the relevance of a predicted endpoint, defining and communicating acceptable model performance characteristics, and providing objective estimates of the reliability and confidence associated with predicted results obtained for specific chemical structures or classes. Notably, the applicability domain of in silico approaches is limited to organic compounds only; normally inorganics and organometallics are outside of the model domain. Whether for internal (e.g., considerations for worker safety) or external use (e.g., for submission to a regulatory authority), these and similar factors are each capable of contributing to a reluctance on the part of users to accept the business risk of introducing in silico-based evaluations which, for example, have the potential to result in an erroneous discovery chemical triaging decision, or could potentially delay a regulatory review.

Presumably, part of the reason for limited detailed guidance is that a consensus within and among agencies, stakeholders, and the risk assessment community as to what constitutes an acceptable in silico data package has yet to emerge. As such, with respect to submissions to a regulatory authority, the impetus to persuade the reviewing agency that prediction data generated using in silico methods is relevant (i.e., reflecting the usefulness of the model for predicting the toxicological endpoint of interest) and reliable (i.e., reflecting the quality of the information used for the assessment), and can be used confidently within stated defined limits to support an overall safety evaluation or hazard assessment, rests with the developer of the method or its user (e.g., the submitting entity).

To address this limitation, an international consortium of experts in the development, use, and application of in silico methods, representing several industry sectors, Federal agencies, research institutes, and academia was organized to develop in silico toxicology (IST) protocols, each focusing on a specialized area of toxicology. In 2018, the consortium published a general framework to outline topics to be addressed in each protocol [8,9]. IST protocols have been published for genetic toxicity [10] and for skin sensitization [11]. Additionally, position papers highlighting points for consideration based on the status of available in silico support have been published for organ toxicity, neurotoxicity, carcinogenicity, and confidence assessment [12–16].

The objectives of IST Protocols are to:

a) increase an understanding of how in silico methods can be used either alone or to supplement hazard assessment and safety evaluation submissions;

b) identify the most common areas of applicability for the particular type of effect addressed in the protocol (e.g., acute toxicity; carcinogenicity; genetic toxicity; neurotoxicity; organ toxicity; skin sensitization);

c) identify several of the most used experimental approaches for assessing toxicity, including assays and endpoints for which in silico models exist or present an opportunity; and.

d) provide recommendations for communicating in silico-derived data, including information related to the relevance and reliability of results, and to the overall level of confidence associated with an evaluation.

It is anticipated that by addressing key considerations associated with these objectives, IST Protocols will foster increased use and acceptance of in silico-derived data in those areas of application for which they are developed.

The present IST protocol centers on in silico-based evidence as applied to support evaluation of acute toxicity. Given the diverse areas of application, scenarios, and purposes for assessing acute toxicity (i.e., questions needing to be addressed for-cause), the intent of the protocol is to provide a generally applicable conceptual framework and not just guidance for satisfying regulatory submission requirements. To develop this framework, emphasis is given to application of in silico methods supporting weight-of-evidence approaches, such as those commonly used in hazard identification and classification. The same principles, which are aimed toward assessing the level of confidence in an evaluation by improving transparency with respect to in silico model development, testing and performance, for communicating data relevance and reliability, and for exploring limitations of an approach, are expected to equally apply in other arenas.

For scenarios requiring a formal writeup (e.g., when the assessment is part of a submission to a regulatory authority), a clear understanding of the purpose and objectives for making the assessment lays the groundwork for presenting material in a rational, well organized, and persuasive fashion, all of which help to facilitate the

review process (i.e., presenting the problem formulation). Considerations for reporting are included following the discussion on relevance, reliability, and confidence.

1.2. Alternative approaches

For many purposes, regulatory agencies continue to require acute toxicity testing in animals, particularly for chemicals where exposure is likely to be significant, such as when the chemical is a primary (active) ingredient in a commercial product. To decrease the number of animals used in acute toxicity studies, advanced study designs have been introduced that, while minimizing the number of animals, maintains the reliability and usefulness of the data. In addition to studies in rodents, there is growing interest in the use of non-mammalian species (e.g., zebrafish, fruit flies, nematodes) as representing whole organism models which can potentially be calibrated to predict the rodent LD50 [17–20]. Other strategies for decreasing animal use in assessing acute toxicity include the use of in vitro assays, tiered testing strategies, and application of in silico methods, often in combination [21–23].

Experiment-based alternative testing paradigms generally consist of sets of in vitro assays. In choosing which assay(s) to conduct, primary consideration is in determining the endpoint(s) most relevant to the scientific objective of the investigation (e.g., evaluation or characterization of acute toxicity) and the specific purpose for conducting the experiment (e.g., compound selection, identifying starting doses for other studies, product safety, hazard identification, setting both threshold and limit exposures). Considerations on human relevance or on the mechanism preserved across species (rat and humans) should also be included.

In settings where in vivo testing in mammals is not mandated, a tiered strategy combining non-testing (i.e., computational approaches, including the use of in silico models) and testing (e.g., in vitro studies, studies in phylogenetically lower species, etc.) approaches may be beneficial for prioritizing activities used to assess acute toxicity. In general, tiered strategies utilize less expensive methods with faster turn-around times in the lower tiers and, when warranted, additional testing successively progresses toward the higher tiers involving more expensive, resource intensive methods. Importantly, testing at a higher tier only occurs when the former tier fails to provide adequate, fit-for-purpose, decision-level data. An excellent example of the tiered approach for assessing acute toxicity as related to chemical defense is that recommended by The National Academies of Sciences, Engineering, and Medicine [20].

As noted above, non-testing approaches, such as in silico predictions, are an important component of a tiered approach. Arguably, it is feasible to develop in silico model(s) for any endpoint generated by a test method at

any level of a tier, provided that a source of adequate reliable experimental data is available or can be assembled. Since in silico methods primarily rely on the availability, quality, and breadth of chemical coverage of data already existing for an assay or study endpoint, predictive models can potentially replace the need to conduct one or more tests, including those normally performed in one of the higher tiers. Similar to when developing a strategy based on a testing approach, the relevance of endpoints modeled using in silico methods and their appropriate use for an intended purpose must be considered. In addition, model performance characteristics, generally determined at the time of model development and testing, together with runtime performance indicators (i.e., those obtained as output when making predictions on structures of interest), must be objectively evaluated, and communicated.

1.3. Areas of application

Knowledge of potential acute toxicity associated with a chemical substance or mixture is vital for protecting society against the harmful effects of chemical exposure. In some arenas, for example where the primary purpose of an evaluation is hazard identification, a weight-of-evidence approach is deemed sufficient in lieu of direct testing in animals. Utilizing this strategy, data that already exist can often be leveraged

to predict potential adverse (i.e., hazardous) effects.

Assessment of acute toxicity is most often needed in the areas of consumer, occupational and environmental safety, and in public health, where activities can generally be classified into one of three broad categories:

a) Product research and development (R&D), which can include screening for active moieties and for triaging compounds for candidate selection, and which may include assessment of metabolites, and of residuals originating from process manufacturing.

b) Product approval or registration, which requires a more thorough assessment of active ingredient(s), metabolites, impurities, and contaminants resulting from manufacturing, storage, and degradation.

c) Transport, storage and handling of product, product intermediates, additives, and chemicals used in manufacturing, where the goal is often to provide data needed to fulfill international chemical registrations, with the intent of informing occupational and environmental hazards that can result from spills, leaks, and other forms of release [24].

These categories can be more granularly subclassified into areas of particular interest at a given point in time along a milestone pathway (e.g., for new product development from initial discovery through development and manufacturing, approval/registration, and culminating at the handling, storage, and transport stages). At each

milestone point, an assessment of potential safety issues provides important information related to activities, processes, operations and other health or environmental concerns. As expert knowledge and statistical models are developed and made available for endpoints relevant to an assessment, evaluations relying in whole or in part on in silico approaches are playing an increasing role, particularly where experiment-derived test data for the chemical of interest are sparse or non-existent.

Table 1 identifies areas of activity where evaluation of acute toxicity commonly applies and where there is an opportunity for in silico models to contribute.

Table 1. Areas of applicability of acute toxicity assessment and examples of uses for in silico methods.

Area	Use	Opportunities for *In Silico*
Product R&D	Scaffold selection	Screening away from off-target acutely toxic compounds; often performed on compounds representing structural scaffolds of potential interest for new product development
	Candidate selection	Choosing from among several molecules within one or more structurally active series with a goal of increasing the probability of technical success of a project by avoiding those having a greater safety liability (i.e., optimizing investment of resources)
	Study design support	Selecting dose ranges, salt forms, and additional endpoints to optimize the design of *in vivo* and *in vitro* studies (e.g., *in silico* results can help identify potential target organs or mechanisms of toxicity, which can then be included as endpoints for investigation)
	Metabolite analysis	Identifying probable test article metabolites and inform on their potential toxicity (e.g., through a bioactivation mechanism)

	Issue resolution	Informing on potential mechanism(s) that could be associated with an observed toxicity; useful for forming hypotheses for testing. Confirmed mechanisms can be screened against to find more suitable candidates
	Scaffold hopping	Identifying alternative structural series which continue to possess desirable properties without the continued presence or degree of undesirable off-target properties
	Weight of evidence	Providing a basis for decisions when little or no data are available that used direct testing methods. Results are also useful for determining whether additional testing is needed
Manufacturing and Process Chemistry	Green chemistry support	Selecting occupationally or environmentally safer starting materials and additives
Occupational Health and Safety	Worker safety	Identifying compounds likely to pose a hazard to workers, e.g., through handling or accidental exposure
	Process control and containment	Informing on the potential need for additional handling precautions
	Exposure limits	Assisting in establishing safe exposure thresholds and limits
	Safety Data Sheet support	Identifying the appropriate GHS classification when *in vivo* and *in vitro* test data are limited or do not exist
Product Safety and Registration	Safety data for registration/approval	Supporting product safety assessment (e.g., by providing the likelihood of involvement of a specific mechanism relating to a study finding)

	Product quality and specifications	Supporting setting of limits for contaminants (e.g., residuals, leachables and extractables, material interaction and degradation products)
	Classification and Labeling	Informing labeling for intended use(s)
Public Health and Safety / Environmental	Classification for product transport	Identifying the appropriate GHS classification, when *in vivo* and *in vitro* test data are limited or do not exist, and assignment of the proper packing group.
	Environmental discharge limits / disposal	Assisting in establishing limits and specifications for restricting environmental pollutants
	Accidental release situations	Informing assessments of risk and strategies for mitigation or countermeasures
Emergency Response	Intended or unintended acute exposure	Informing assessments of risk and strategies for mitigation or countermeasures
Military	Protection of military personnel	Identifying hazards and informing assessments of risk and strategies for mitigation or countermeasures (i.e., tactical preparedness)

1.4. Regulatory landscape

Many arenas rely on an evaluation of acute systemic toxicity to inform hazard identification and to meet regulatory requirements. In cases involving worker, consumer, and public health and safety, government agencies have established regulations and published guidance on how to comply with requirements [25]. Generally, differences in requirements across agencies reflect the scope and priorities, as established legislatively. An area that tends to be utilized is acute toxicity assessments made using a Weight-of-Evidence (WoE) strategy, and which may include results based on in silico modeling [22,26]. ECHA reports regularly on the usage of various test methods to support registrations and recently reported that QSAR methods were utilized

to fulfill acute toxicity information requirements between 1 and 2 % of the time in 2019 (based on 94,551 acute toxicity records in 2019) [27]. Strickland et al. recently published an article reviewing the status of acute systemic toxicity requirements and data uses by U.S. regulatory agencies [25].

The Organization for Economic Cooperation and Development (OECD) leads international efforts to harmonize regulatory testing approaches and has validated several testing approaches for conducting acute systemic toxicity testing in animals. Important considerations include a determination that the animal model is relevant to humans and that the chosen route of exposure reflects situation(s) most likely to be encountered in a real-world setting, e.g., during intended product use, in an accidental exposure scenario, etc.

Current testing strategies across countries requiring international chemical testing registrations (e.g., EU and Korea REACH, China MEP registrations, etc.) include an effort to reduce reliance on animal testing, primarily by replacing required test endpoints with in vitro and/or in silico alternatives. For example, in vitro skin and eye irritation studies have recently been accepted by the EU for registrations [28]. However, in vivo studies are still required in other jurisdictions, making it difficult to fully eliminate animal testing. Similarly, one of the better known and standardized in vitro methods used as a surrogate endpoint for acute toxicity is basal cytotoxicity (i.e., basal cell lethality), that has gained wider regulatory acceptance. A testing protocol for the ICCVAM validated 3 T3 Neutral Red Uptake (NRU) cytotoxicity assay has been published and is accepted or can be proposed for some purposes, such as setting starting doses for in vivo studies [29–31].

Because submission requirements vary among regulatory authorities and other federal agencies that use the information, and are periodically updated, often on a planned publication release schedule, it is always best to consult the most recent version of a regulatory or guidance document, which is available for download from the agency's website. In addition to regulatory documents, most agency websites provide a significant amount of other useful information, including links to scientific references, technical documents and protocols, checklists and templates used for submissions, descriptions of collaborative research activities, informatics projects and tools, public data sources, and sources of other relevant information, which may include links to other agencies. It is highly recommended that users become familiar with agency publishing schedules and consider joining mailing lists to receive automatic notifications of important updates. As noted earlier, challenges associated with more fully utilizing in silico approaches by a regulatory authority include ascertaining the level of uncertainty with respect to predictive accuracy and to then effectively communicate results in a way that instills confidence that an evaluation is appropriate, fit for purpose, and reaches a justifiable conclusion based on a reasoned interpretation of results. Moreover, by identifying, acknowledging, and communicating limitations of an approach, for example by delineating reasoned boundaries with respect to how in silico results are applied and

interpreted for the stated purpose, agency reviewers are provided with supporting information needed to enable an independent evaluation of overall conclusions of an assessment.

1.5. Hazard assessment framework

Acute toxicity encompasses a broad set of adverse effects which might occur following acute exposure to a chemical agent. Fig. 1 presents a Hazard Assessment Framework (HAF) relating to assessment of acute toxicity. While the HAF serves to highlight the complexity of acute toxicologic responses, it also provides a high-level view of the many data streams available for developing in silico methods that might feed into a WoE approach which might incorporate knowledge from a combination of in vivo, in vitro, and mechanistic studies.

A primary consideration for any assessment of acute toxicity includes an objective determination of those factors which are likely to influence the probability of observing toxicity. These factors include exposure scenario(s) that dictate routes, concentrations and times of exposure, the physical form and/or composition of various formulations or mixtures, and properties that influence absorption, distribution, metabolism, and excretion (ADME). In other words, it is generally not sufficient to only assess the potential innate toxicity of a chemical agent for a particular site of action (e.g., organ or tissue) based, for example, on structural similarity to other compounds known to cause toxicity at that site, or to a known mechanism of action (e.g., interaction with a particular molecular target). Rather, exposure considerations are likewise a key component of an assessment. Target organs of toxicity associated with acute exposures are most often either those that come into direct contact with a chemical (e.g., skin and eyes, gastrointestinal tract, respiratory tract), where general caustic or cytotoxic effects can occur, or those involved in the maintenance of homeostatic or rapid signaling responses in vital organs (e.g., electrolyte and fluid balance, ion channel activity in the nervous and cardiovascular systems, oxygen uptake and utilization, and energy production). In vivo and in vitro assays targeting a wide array of specific mechanisms known to be associated with these types of critical processes have been developed for assessing whether a chemical can potentially act through that mechanism to produce toxicity [1,20, 32–36].

In addition to experiment-based testing, data from previously conducted studies performed on structurally similar compounds or on chemicals belonging to the same chemical class are often available and can be leveraged directly or indirectly, for example as part of a training set for development of in silico (Q)SAR prediction models, to support a WoE assessment approach (Fig. 1).

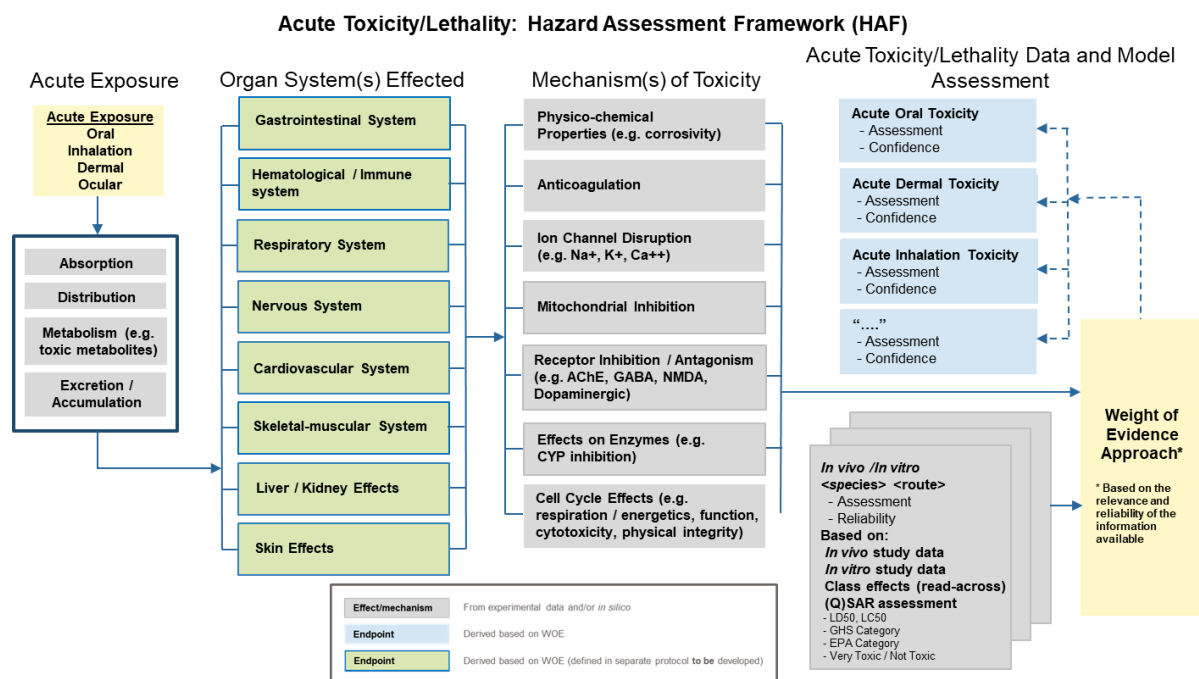**Acute Toxicity/Lethality: Hazard Assessment Framework (HAF)**

Figure 1. Hazard Assessment Framework (HAF) supporting assessment of acute toxicity

Given the diversity of mechanisms, organ/tissue systems, and mitigating or accentuating factors that can interact with one another to produce toxicity, and considering the uncertainty associated with accurately predicting adverse outcomes in humans from animal studies, a "one-size-fits-all" approach capable of addressing the central question (s) for all purposes is not feasible. For these reasons, the HAF shown in Fig. 1 is intended to be a helpful guide only and not a prescribed path to be followed when making an assessment. In most cases, some knowledge of a compound's toxicologic profile, based on its chemical class, intended target (as it applies to pharmaceutical or pesticide products), physical properties, and general screening results (lab and/or in silicobased) will be available. Information pertaining to each of these is helpful for making informed decisions as to which in vivo, in vitro, and/or in silico tests will be most helpful for evaluating the acute toxicity liability.

1.6. Globally Harmonized system classification

A WoE approach is often used to classify chemicals in accordance with the Globally Harmonized System of Classification and Labeling of Chemicals (GHS) [3]. This classification system was developed chiefly in response to the 1992 United Nations Conference on Environment and Development (UNCED) mandate for an internationally harmonized chemical classification and labeling system that could be used to identify and rank

the severity of physical, health, and environmental hazards associated with a chemical [37]. The UNCED mandate was endorsed by the United Nations General Assembly to "provide the foundation for all countries to develop comprehensive national programs to ensure the safe use of chemicals" [38,39]. Accordingly, categories for GHS health and environmental endpoints, including acute toxicity have been developed and adopted for context, as appropriate, by agencies worldwide. In the U.S., these agencies include OSHA (Occupational Safety and Health Administration), EPA (Environmental Protection Agency), and DoT (Department of Transportation) (refer to agency websites for further information on the status of alignment and adoption for individual agencies). Resources are available which provide detailed guidance on implementing and communicating GHS classification information [3,40].

Table 2, adopted from the GHS "Purple Book" [3], shows categories for acute toxicity, where exposures are based on measured or estimated (e.g., predicted) values of the rodent LD50. For each hazard class, chemicals are labeled according to the type(s) of hazard they represent and are further categorized by assigning a numerical ranking ranging from 1 to 5, with a lower number signifying greater severity. For most purposes, a limit dose of 2000 mg/kg is used for testing, since compounds having an LD50 of>2000 mg/kg are generally considered non-toxic. The "Toxic Substance" label is defined as "acutely toxic substances that may be fatal or toxic if inhaled, ingested, or absorbed through the skin". Table 3.

Table 2. GHS Classification Categories for Acute Toxicity (from UN, 2021)

Acute Toxicity	Cat. 1 LD50	Cat. 2 LD50	Cat. 3 LD50	Cat. 4 LD50	Cat. 5
Oral (mg/kg)	≤5	>5 ≤50	>50 ≤300	>300 ≤2000	Criteria: • Anticipated oral LD50 between 2000 and 5000 mg/kg
Dermal (mg/kg)	≤50	>50 ≤200	>200 ≤1000	>1000 ≤2000	• Indication of significant effects in humans*
Gases (ppm)	≤100	>100 ≤500	>500 ≤2500	>2500 ≤5000	• Any mortality at class 4*
Vapors (mg/L)	≤0.5	>0.5 ≤2.0	>2 ≤10	>10 ≤20	• Significant clinical signs at class 4*
Dust and Mists (mg/L)	≤0.05	>0.05 ≤0.5	>0.5 ≤1.0	>1.0 ≤5	• Indications from other studies*

*If assignment to a more hazardous class is not warranted

Table 3. Standardized *In Vivo* Approaches for Assessing Acute Toxicity in Animals

Route	OECD Test Guideline	Comments
Oral	Fixed Dose Procedure (OECD 420) Acute Toxic Class method (OECD 423) Up-and-Down Procedure (OECD 425)	These guidelines. The OECD 420, 423 and 425 test guidelines were devised to supplant OECD 401 with methods that utilize predetermined doses, reduce animal usage where the initial dose level is based on a small range-finding study, cytotoxicity screens, and/or preexisting data. Following dosing (typically by oral gavage), animals (generally rats) are monitored for overt toxicological signs including death (Acute Toxic Class or Up-and-Down Method) and "evident toxicity" (Fixed Dose Procedure). Note that the limit dose or recommended maximum dose for the acute oral toxicity tests OECD 420, 423 and 425 is generally 2000 mg/kg. - The Fixed Dose Procedure (OECD 420) includes evident signs of toxicity as indicative of acute oral toxicity and does not solely rely on death as an endpoint. It is generally recommended and recognized as the most humane test method and is believed to utilize the least number of animals. - The Acute Toxic Class method and Fixed Dose method (OECD 423 and 425) utilize a stepwise assessment that results in GHS acute oral toxicity classification. - The Up-and-Down method (OECD 425) can be used when an LD_{50}, rather than solely the GHS category, is required. Acute oral toxicity testing can be waived according to many regulatory agencies if the test material is corrosive.
Dermal	Acute Dermal Toxicity procedure (OECD 402)	OECD 402 involves a single uniform application of test article to ≥10% of the animal's body surface area. The rat is recommended, but rabbit or guinea pig have been used. Fur should be removed ~24 hours prior to application of the test article. The test substance can be applied in solution, where any solid should be moistened with water or an appropriate vehicle and held in contact with the skin using a porous gauze dressing and non-irritating tape throughout a 24-hour exposure period. Animals are observed periodically for clinical signs and toxicity throughout the exposure period and afterward for a total of 14 days. The limit dose for the OECD 402 study is 2000 mg/kg. Acute dermal toxicity testing can be waived

		according to many regulatory agencies if the oral LD50 exceeds the limit dose (2000 mg/kg) as well as if the material is corrosive.
Inhalation	Acute Inhalation Toxicity test (OECD 403) Acute Toxic Class method (OECD 436) Fixed Concentration Procedure (OECD 433)	The Acute Inhalation Toxicity test is performed according to OECD 403 or OECD TG 436. Briefly, groups of five rats are exposed via nose-only (preferred) or whole body to a uniform airborne concentration(s) of the test article for a typical duration of 4 hours. The animals are subsequently monitored for 14 days to determine the lethal concentration in 50% of the animals (LC_{50}). The animals are exposed to the test article as a gas, vapor, aerosol, or a mixture of several phases which is dependent on physical/chemical properties of the test article and its typical use. The limit concentrations for OECD 436 are aligned with GHS classifications and are 20,000 ppm/4h for gases, 20 mg/L/4h for vapors and 5 mg/L/4h for aerosols. OECD 433 (Fixed Concentration Procedure) is a reduction/refinement to TG 403 which only utilizes one sex of rat (females) and includes the endpoint of evident signs of toxicity as a measure of acute inhalation toxicity (does not solely rely on death as an endpoint).

Assignment of chemicals to GHS categories is an area where the application of in silico methods for estimating the rodent LD50 is being increasingly relied on to support WoE hazard assessment in the regulatory arena. As an example, Bercu and coworkers discussed the potential for reliable and broad application of acute oral toxicity (AOT) in silico models across different industrial sectors [21]. The efforts being made to validate and demonstrate the utility of QSAR models for the purpose of GHS classification, together with agency expectations for documenting model development, use, and performance provide a general framework with which to discuss applicability and relevance, reliability, confidence, performance, limitations, and reporting considerations that essentially apply to all areas in which in silico technologies may be employed in a primary or supporting capacity [41].

2. Experimental data and in silico models

2.1. Background

Evaluations of acute toxicity are based on data from in vivo animal studies, in vitro experiments, expert review of existing knowledge and experience, and/or in silico methods, and may be a combination of these approaches. In vivo acute toxicity studies often follow well-established study design protocols as well as guidance from regulatory authorities and/or international organizations such as OECD. Provided it is conducted using an appropriate animal species, an in vivo study is considered scientifically relevant and would not normally require justification beyond species selection and doses (exposures) tested. In contrast, the plethora of potential assay endpoints and experimental designs available for in vitro studies makes assay selection an important consideration in terms of relevance of assay endpoint(s) and the experimental design to support the intended purpose of an evaluation. In addition, assay robustness (i.e., repeatability under standardized conditions) and the reliability and interpretability of the data it generates are important factors for justifying selection of an assay system.

Klimisch et al. [42] published a broad, widely used scoring system for categorizing the reliability (defined as inherent quality) of experimental data, in which a lower score signifies better data quality. Reliability scores are assigned as RS1 (Reliable without restriction); RS2 (Reliable with restriction); RS3 (Not reliable); and RS4 (Not assignable). Myatt et al. [9] provides a thorough description of this scoring system and how it might also be used to categorize the reliability of results based on in silico methods.

Considering that data generated by experiment-based tests provide the raw input for training and testing in silico models, whether statistically- based or expert system-based, objective evaluations of endpoint relevance, assay robustness, and data reliability take on even greater significance. For example, introduction of inaccurate, imprecise, or contradictory data into an in silico model can influence the reliability of predictions across the model's prediction space. Since general models, which are intended to provide broad coverage of chemical space, are most often used over extended periods of time (sometimes for several years before being updated), they are likely to be used to support numerous toxicity evaluations across a diverse set of chemical structures. Prediction errors resulting from inaccurate training data may not be detected but will be a persistent characteristic affecting model performance and extending across all evaluations for which the model is used.

2.2. In vivo studies

2.2.1. Experimental approaches

In most sectors, direct testing in animals continues to be the mainstay for assessing systemic acute toxicity. The classic design involves administering a single dose of test substance (chemical or mixture), with the clinical signs, body weight, occurrence of severe toxicity, and number of deaths being monitored over the subsequent 14-day period. Depending on the results from the initial dose, a higher or lower dose may be administered to an additional group of animals. The primary study endpoint of an acute toxicity study is the dose resulting in mortality or morbidity to fifty percent of animals within 14-days of a single exposure to the test substance (i.e., the LD50). The rodent LD50 is the most widely accepted standard measure for gauging acute mammalian toxicity at the whole animal level and is often cited in regulatory guidance. As such, the need to justify the relevance of the LD50 endpoint is rare. Even when considering alternative methods for acute toxicity, an estimate of the rodent LD50 based on in silico evaluation or read-across to structurally similar compounds often provides the desired result [43].

The OECD has validated several approaches to guide in vivo acute toxicity testing. For example, routes of exposure should be relevant to those expected to be encountered by humans (e.g., dermal exposure, inhalation exposure, oral exposure). To accommodate the differences in exposure routes, OECD has developed acute toxicity testing guidelines for the oral, dermal, and inhalation routes [44].

The general goal of the OECD-driven acute studies is to ascertain the LD50. However, these guidelines as well as UN GHS also recommend a limit dose of 2000 mg/kg and discourage testing at higher doses for animal welfare reasons, unless there is a strong likelihood that results of such a test have a direct relevance for protecting human or animal health or the environment [3,45–47]. In general, if no toxicity or no significant clinical signs occur at the limit dose, the study can be deemed complete with the lethal dose concluded to be greater than the limit dose (e.g., LD50 > 2000 mg/kg). Moreover, it is important to note that more is gained from these studies than the dose at which mortality occurs. Clinical signs and symptoms may provide evidence of tolerance, identify organ system involvement at a dose-limiting toxicity, and can suggest mechanisms that may lead to a greater understanding of the acute effects of the test substance.

Concurrent animal controls (e.g., treated with vehicle) are generally not warranted for these studies. Additionally, for compounds deemed corrosive, in vivo acute toxicity studies can be waived based on animal welfare and ethical considerations.

2.2.2. In silico approaches to predicting in vivo outcomes

As noted earlier, LD50 is the accepted standard and in silico methods for predicting LD50 or categories derived from LD50 values (such as GHS categories) are prime candidates for model development. In silico models for acute in vivo lethality have been summarized and compared in a number of publications [53–59]. One critical factor in the development of such models is the availability of a sufficient quantity of high quality in vivo acute toxicity data. Table 4 provides a summary of some sources of acute toxicity data to support in silico model development and readacross predictions. The information in the table is derived from a number of sources: the Registry of Toxic Effects of Chemical Substances (RTECS©), the US Food and Drug Administration's Center for Food Safety and Applied Nutrition (CFSAN) Priority Based Assessment of Food Additives (PAFA) database, the European Chemicals Agency (ECHA) ChemProp database, the European Union's Joint Research Centre (JRC) AcutoxBase, the National Library of Medicine (NLM) Hazardous Substances Data Bank (HSDB), the Organization for Economic Co-operation and Development (OECD) eChemPortal, and TEST (NLM ChemIDPlus) [60–63]. Table 4 illustrates the different number of chemicals with test results based on species and route of administration. It should be noted that there will be overlapping numbers of chemicals across different sources. Most model development to date has focused on rat oral lethality because of the large numbers of chemicals with LD50 values in the public domain available for modelling. Other endpoints have fewer numbers of chemicals which can make modelling more challenging. These models have been developed to predict a number of endpoints related to acute lethality. As noted above, these include LD50, hazard classifications (such as the GHS or EPA classifications), non-toxic classification (often defined as LD50 > 2000 mg/kg), and very toxic classification (often defined as LD50 < 50 mg/kg) [64].

Table 4. Summary of different sources of *in vivo* acute toxicity data

Source	Study type	Endpoint	Number of unique chemicals
RTECS	Rat oral	LD50	16,499
RTECS	Rat dermal	LD50	1,267
RTECS	Rat inhalation	LC50	1,718
RTECS	Mouse oral	LD50	34,522
RTECS	Mouse dermal	LD50	264
RTECS	Mouse inhalation	LC50	915
RTECS	Rabbit dermal	LD50	5,321
CFSAN/PAFA	Rat oral	LD50	949
CFSAN/PAFA	Mouse oral	LD50	366
ECHA (ChemProp)	Rat oral	LD50	2,136
JRC AcutoxBase	Rat oral	LD50	138
NLM HSDB	Rat oral	LD50	2,205
OECD (eChemPortal)	Rat oral	LD50	2,290
PAI (NICEATM)	Rat oral	LD50	293
TEST (NLM ChemIDplus)	Rat oral	LD50	12,974

In silico methodologies include statistical or Quantitative Structure- Activity Relationship (QSAR) models and a limited number of expert rule-based models. Moreover, local models have been developed that focus on a specific class of chemicals, such as N-nitroso compounds [65], sulfur mustard derivatives [66], aromatic chemicals [67], organophosphorus pesticides [68], and possible mechanisms of nonspecific action [69], alongside global models covering chemicals within the applicability domain [70–72]. The automated read-across workflow implemented in the OECD QSAR Toolbox is a decision-scheme that predicts acute oral toxicity using a mechanism-based category approach and considers metabolic similarity [43]. Table 5 summarizes the algorithms that have been used to develop models for the prediction of acute oral toxicity.

Table 5. A variety of algorithms (and structural and physico-chemical descriptors) can be used to develop models for the prediction of acute oral toxicity.

Algorithms	References
Random Forests	[73-78]
Artificial Neural Networks	[74,79,80]
Deep Learning	[78,81-83]
Local Lazy Learning	[84]
k-Nearest Neighbors	[71-74,78,85]
Support Vector Machines	[74,77]
Arithmetic Mean Toxicity	[86]
Partial Logistic Regression	[87]
Partial Least Squares Regression	[78,87]
Multi-Descriptor Read Across	[88]
Clustering-based QSAR model	[89]
Multiple Linear Regression	[78]
Global, Adjusted Locally According to Similarity	[90]
Decision Trees	[78]
Expert rule-based methodology	[21]
Read-Across Structure Activity Relationships	[43,91]
Naïve Bayesian	[77]

A number of publications have highlighted the importance of providing a clear domain of applicability assessment for global acute toxicity models applied to new chemicals [92,93]. In addition, consensus modelling approaches have been successfully adopted in a number of publications highlighting the power of combining models [94–96]. The results from consensus in silico models were independently assessed using public data as part of a 2018 workshop [97] and the conclusions from the organizing committee was that the results "… were equivalent to the ability of the rat oral LD50 data to predict itself" [64]. In other words, combined results from different models (defined under the general in silico framework as having a reliability score RS4 or above) are fit-for-purpose for predicting rat oral lethality. This was also the conclusion from Bercu and co-workers when such models were applied to predominantly proprietary data [21].

It has also been shown that read-across can be used to predict acute in vivo lethality using a combination of public and proprietary data [7, 53].

When assessing in silico performance, it is important to consider the performance of the in vivo test itself. Hoffmann et al. [98] showed that for a limited number of chemicals with multiple test results, 54 % would fall into the same GHS category and 44 % would fall within adjacent categories. Karmaus [61] performed a similar analysis on a much larger collection and found that, for chemicals tested more than once, 74 % would fall in the same GHS category.

Graham et al. [27] evaluated a set of 371 internal compounds from an historical acute toxicity LD50 database of pharmaceutical intermediates and active pharmaceutical ingredients. Using two statisticallybased models, they found that 77–95 % of predictions fell within one GHS category of the experimentally assigned category. Predictivity was generally better for compounds with experimental LD50 > 300 mg/kg (i.e., GHS categories 4, 5, and Not Classified (LD50 > limit dose, generally 2000 mg/kg)). Bercu and co-workers assessed the application of statistical-based and expert rule-based models to predict GHS categories. It was shown that the individual models were able to predict either the correct category or a more conservative category for over 90 % of the chemicals. A consensus prediction based on both methodologies was also evaluated and had the highest score for correct or more conservative [21].

Collectively, these evaluations of in silico performance indicate that predictions obtained using in silico approaches are now well-suited for reliably supporting assessment of LD50-based acute toxicity for the purpose of GHS classification. This is further underscored by considering that the inherent variability of results from in vivo studies, in combination with the sharp cutoff values used to define GHS categories, makes experimentally

determined categories no more accurate than to within one category, particularly when the LD50 is close to a category boundary.

2.3. In vitro studies

2.3.1. Experimental approaches

The biochemical and biophysical diversity of mechanisms and associated adverse outcome pathways (AOPs) that can elicit an acute toxicologic response are multifarious. While key initiating events for some toxicities are readily understood in terms of known physical or biochemical mechanisms (e.g., direct-acting caustic or detergent action on tissue, inhibition of critical enzymes, cell receptor proteins, ion channels, etc.), many others remain to be identified. Moreover, compounds acting through multiple pathways are challenging because of the greater complexity of interactions involved, many of which are likely unknown.

A consequence of these considerations is that no current set of in vitro endpoints is broad enough to provide complete coverage of all potential mechanisms or interacting factors capable of influencing the expression and/or magnitude of an acute toxicologic response at the whole organism level. For example, in vitro tests generally do not account for physiological processes capable of modulating toxicity in intact organisms. These processes include those that underpin differences in exposure at a site of action in addition to those involving redundant or compensatory pathways which can obscure effects occurring through other mechanism(s). In these scenarios, toxicity can be exacerbated (e.g., locally within an organ/tissue through bioaccumulation or bioactivation) or mitigated (e.g., through metabolic detoxification, limited distribution to a site of action, active cell efflux, or again through compensatory physiologic pathways).

The inability of in vitro systems to fully recapitulate many of the important and simultaneously acting biochemical and physiologic processes operating at the organismal level limits their usefulness for extrapolating to whole animal toxicity. Nevertheless, in vitro tests are often useful for directly evaluating interactions that involve a specific mechanism or mode of action and investigations incorporating targeted biochemical endpoints can be useful for generating mechanistic hypotheses and for identifying potential pathways leading to the toxicologic effect(s). Targeted mechanism-based assays can also be used to rule out a potential mechanism as the basis for an observed toxicity, or to provide reasonable confidence that a particular toxicity will not be observed.

Moreover, in vitro assay conditions can mimic processes operating in vivo (e.g., introduction of key metabolic enzyme preparations into the test system to generate metabolites). In vitro studies can also be useful for providing a rationale for establishing exposure boundary limits, and often provide insights into potential chemical class effects. Lastly, organ specific toxicities can be assessed and measured using more complex tissue-chip systems that recapitulate critical physiological functions sensitive to acutely toxic chemicals. These systems better represent human physiology than traditional in vitro systems by incorporating primary cell types with three-dimensional structure and mechanical stress [99].

Provided assay endpoint(s) have been validated scientifically as having relevance to acute toxicity and demonstrated assay reliability and robustness, in vitro tests can provide important information. Addressing assay relevance and reliability are two key requirements for establishing the state of "test readiness" for regulatory acceptance of in vitro data [100,101].

Although results from in vitro assays are often not adequately calibrated to in vivo endpoints, such as the LD50, this continues to be a goal. Accordingly, results from in vitro tests are generally positioned to serve in a supporting role for assessments of acute toxicity, for example to:

• Screen and/or triage compounds for further study (e.g., product candidate selection; environmental testing prioritization).

• Aid in the design of in vivo studies (e.g., dose selection; studies investigating putative mechanisms or a most probable set of expected outcomes).

• Establish (verify or refute) a cause-effect relationship based on a particular mechanism (key initiating event or toxicologic pathway (i.e., Adverse Outcome Pathway).

• Provide data for a WoE-based assessment.

• Elucidate novel mechanisms/pathways of toxicity.

The advantages of in vitro testing in terms of cost, time, and capacity, together with their ability to provide mechanistic insights, make in vitro assays a powerful approach for conserving resources and for further reducing reliance on animal studies. Their ability to support the rapid evaluation of large numbers of compounds (as compared to in vivo testing), additionally provides the opportunity to generate data across a wide range of structurally diverse chemical series. Moreover, the large amount of data generated across structurally diverse compounds provides a basis for developing robust in silico approaches that can be applied globally to novel chemical structures to predict probable responses in the modelled assay.

In vitro studies can generally be classified into non-mechanistic assays based on endpoints expected to broadly apply across tissues (e.g., cell lethality and necrotizing corrosiveness) and mechanistic assays designed to assess

specific molecular interactions which are known to be associated with certain toxicologic sequela (e.g., ion-channel inhibition and disruption of cellular energetics).

Non-mechanistic assays are useful for gauging the potency of a chemical with respect to dose-limiting toxicity, whether it occurs through a known or unknown mechanism. This information is often useful for assigning a compound to a potency class or to set an upper limit of exposure, e.g., as when choosing an upper dose for additional studies. As such, non-mechanistic assays are widely employed for chemical screening and classification purposes when detailed knowledge of the toxicologic mechanism is less important than estimating the exposure likely to result in an adverse outcome of any type.

Conversely, mechanism-based assays provide insight on how a chemical might generate a toxicologic response, so are helpful for developing and testing hypotheses in addition to screening, classification, and estimation of potency for interactions occurring through that mechanism, whether it is dose-limiting or not. For example, it is reasonable to anticipate that a novel compound within a chemical series is likely to interact via the same mechanism that has been previously established for other members of the class. In this case, a targeted mechanismbased assay is ideal for testing the mechanistic hypothesis.

Another way to view the difference between non-mechanistic and mechanism-based assays is that non-mechanistic assays pool an entire set of potential mechanisms, known and unknown, into a single measurable endpoint (e.g., cell death), whereas mechanism-based assays parse potential adverse events into known discrete mechanisms which can then each be tested independently (e.g., aerobic oxygen utilization, uncoupling of oxidative phosphorylation, inhibition of hERG ion channel function).

Selecting the most appropriate in vitro assay(s) to run as part of a toxicology assessment should be based on a reasoned testing approach designed to address the primary goal of the assessment and not simply on which assays are available. Two strategies are employed when determining which test(s) to conduct for an intended purpose:

1) When existing data are available for structurally similar compounds known to operate through a particular mechanism (i.e., key initiating event or pathway), an in vitro assay (or in silico counterpart) for that endpoint is used to determine whether the test compound is likely to engage that mechanism. In this scenario, the number of endpoints evaluated are often limited to those previously established as being relevant to and dose-limiting for the chemical class.

2) When data are sparse for structurally similar compounds, potential mechanisms of toxicity are not appreciably known, and a more general testing approach is used. This approach may include assay(s) for which the endpoint is not dependent on a single mechanism (i.e., a nonmechanistic assay such as cytolethality), or that consist of a

battery of complementary assays with endpoints known to be relevant to the primary (apical) endpoint. Assays chosen for this second approach are often hierarchically organized into a tiered testing strategy where the lowest tiers generally include one or more non-mechanistic assays for making a preliminary assessment of the toxicologic landscape. Results from the lower tiered tests can then be used to determine whether and which additional higher tiered, targeted mechanism-based assays are warranted.

Advances in the biomedical and related sciences over the past few decades have led to a dramatic increase in the number of hitherto unknown proteins and their associated function(s). This knowledge has, in turn, broadened our understanding of key triggering mechanisms operating within Adverse Outcome Pathways (AOPs) that lead to toxicity. One of the consequences of this explosion of available information is that the number of in vitro assays available for screening/testing and the number of vendors offering such services continues to grow each year, making it impractical to provide a complete listing of relevant assays and mechanisms within most IST protocols.

In addition to the HAF shown in Fig. 1, Table 6 includes examples of in vitro test endpoints/targets having a well-established relationship to acute toxicity. A recent workshop on alternative acute toxicity testing [1] also identified a number of relevant in vitro endpoints.

Table 6. *In vitro* endpoints commonly evaluated for predicting acute toxicity

Test Category/Topic	Example	Rationale (for acute toxicity)
1. General		
Physico-chemical Properties	cLogP, pKa, H-bond donors/acceptors, MW, PSA	Influence biophysical effects: 1) can relate directly to toxicity (e.g., corrosives); 2) can modify expression of toxicity (e.g., through effects on ADME properties) Are often included as chemical descriptors used to develop *in silico* (Q)SAR models
	Michael-Acceptor	Chemical reactivity (e.g., adduct formation)
Cytotoxicity/Cytolethality	Non-specific (i.e., basal) cell	Toxicity resulting in cell death generally (i.e., without regard to effects that may be limited to specific tissues or cell types)

	Specific cell lineages	Specificity of toxicity directed toward certain tissues/organs (e.g., cardiomyocytes, hematopoietic cells)
Cell Growth and Proliferation	Non-specific or specific for certain cell lineages	Effects related to cell stasis without necessarily causing death (e.g., CDK inhibition)
Cell Energetics	ATP depletion	Ability to provide energy for cellular processes
	Mitochondrial function	Ability to conduct oxidative (aerobic) metabolism
	Ox-Phos uncoupling	Ability to capture and store energy from electron transport (oxidative metabolism)
Other Cell Functions	Protein synthesis	
	DNA, RNA synthesis	
	Cell-cell signaling	Ability of cells to interact with their environment (e.g., ligand-receptor interaction)
	Secretory function	Production of major secretory substances (e.g., hormones)
Cell/Tissue Morphology	Vacuolation; Accumulation	Detection of morphologic abnormalities
	Membrane Integrity	Effects on plasma membrane function (e.g., cell homeostasis involving compartmentalization, maintenance of gradients, etc.)
Test Article Metabolism	GSH Depletion and/or Adduct Formation	Effects on normal detoxification pathways.
2. Targeted Mechanisms		
Specific Enzymes	Acetylcholinesterase	Cholinergic effects
	Cytochrome P450 (CYP) enzymes	Drug/toxicant metabolism
Receptors and Ion Channels	**hERG** (hERG): Potassium voltage-gated channel subfamily H member 2	Cardiac function
	Cav1.2 (CACNA1C): Voltage-dependent L-type calcium channel subunit alpha-1C	Cardiac and neurologic function

	Nav1.5 (SCN5A): Sodium channel protein type 5 subunit alpha	Cardiac function
	Kv4.3 (KCND3): Potassium voltage-gated channel subfamily D member 3	Cardiac function
	KCNQ1: Potassium voltage-gated channel subfamily KQT member 1	Cardiac function
	KCNE1: Potassium voltage-gated channel subfamily E member 1	Cardiac function
	Kir2.1 (KCNJ2): Inward rectifier potassium channel 2	Cardiac function
	5HT2B: 5-hydroxytryptamine receptor 2B	CNS function
	B-AR, PDE	
	iPSC-CM	
	Purkinjie fiber assay	
Nervous System	GABA Receptor(s)	CNS function (GABA receptors are the predominant inhibitory neurotransmitter receptors)
	GABA Benzodiazepine Site	CNS function
	NMDA Receptor	CNS function (glutamate receptors are the predominant excitatory neurotransmitter receptors)
	D2 Receptor (DRD2): D(2) dopamine receptor	CNS function

While some mechanisms act non-selectively across all tissues (e.g., oxidative phosphorylation), others act selectively within organ systems that are of particular concern for acute toxicity (e.g., ion channels in cardiovascular and nervous tissue, detoxification systems in hepatic tissue, coagulation in the hematologic system). However, while targeted endpoints may not be a high priority in a general assessment of toxicity,

pharmaco-toxicologic interactions can be associated with undesirable off-target effects and can be screened against during early assessment phases, e.g., when prioritizing compounds for further development.

As might be expected, off-target pharmacologic mechanisms of obvious concern for acute toxicity involve critical organ systems where disruption of normal function, such as those mediating rapid response pathways (e.g., electrochemical signaling, oxygen utilization, energy production) occur quickly and can have life-threatening consequences.

The choice of which in vitro assay(s) to include in an evaluation of acute toxicity depends on the reason for performing the assessment. Arguably, the primary objective is to provide the information deemed adequate to arrive at a reasoned and defensible decision, whether it is for determining which compound to move forward in development, whether the compound is safe enough to be approved for a particular purpose, or how the compound should be classified to ensure proper handling, transport and storage.

2.3.2. In vitro cytolethality assays

While it is outside the scope of the IST Protocol to present detailed information on the various available in vitro tests that can be used for assessing acute toxicity, a brief description of the cytolethality test is warranted due to its widespread use and the availability of standardized, validated protocols.

Cytotoxicity or cytolethality assays measure the intrinsic ability of chemical exposure to result in cell death. Many test systems, including those based on primary cells collected from specific tissues (e.g., hepatocytes from liver, hematopoietic progenitors from bone marrow) and those based on cell lines representing various cell types (e.g., HepG2 hepatocytes, 3 T3 fibroblasts, L6 myocytes), are available for use in these assays. Moreover, in addition to simply quantifying cell death, a number of other endpoints for measuring cell integrity or function, such as membrane leakage, mitochondrial function, and nuclear staining characteristics are available using high-content cytometry methods.

Cytolethality assays are often used in early discovery and development to prioritize compounds and for setting starting doses for in vivo studies. For example, based on validation results for the mouse fibroblast 3 T3 cell line (or human epidermal keratinocytes (NHK) primary cells) test system utilizing a Neutral Red Uptake (NRU) assay [102] which compared the in vitro results to in vivo post-mortem LC50 values for the same set of 72 test substances, the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicologic Methods (NICEATM) promulgated a recommendation for using the assay to aid in setting starting

doses for in vivo studies [103], where it is estimated to reduce animal use for each study by as much as 50 % [104, 105]. OECD published a guideline in 2010 supporting this use [106].

As described in the ECHA's guidance on Information Requirements and Chemical Safety Assessment, the 3 T3 Neutral Red Uptake cytotoxicity assay could be used within a WoE approach to adapt the standard information requirements [29]. The NRU assay is sensitive to hazardous substances acting through general mechanisms of toxicity common to most cell types (basal cytotoxicity) and it well predicts substances with low acute oral toxicity (i.e., those which are not to be classified for acute toxicity); it has a high false positive rate and the interpretation of the negative results should account for the lack of metabolic competence of the 3 T3 cell line and difficulty in capturing specific mechanisms of action in relation to interaction with specific molecular targetsin certain tissues [29].

A recent analysis by Prieto and co-workers noted that general cytotoxicity is a determining factor of acute systemic toxicity and that the majority of the studied chemicals leading to acute lethal toxicity act via some general mechanisms of toxicity rather than organ-specific pathways [34]. It was then noted that the most frequent targets are the nervous and the cardiovascular systems.

2.3.3. In silico approaches to predicting in vitro endpoints

Just as the data collected from in vivo studies enable developing in silico predictive models, it is feasible to replace direct testing performed in vitro with in silico predictions for the same endpoints. Whether this can be realized in practice for a selected endpoint depends on the availability of high-quality experimental data relevant to the chemical space of interest and on model performance.

An effect which occurs as the result of a well-recognized biological or chemical mechanism, such as those generally recognized to be involved in genetic toxicity, genotoxic carcinogenicity, skin sensitization, and skin irritation, are often successfully predicted by in silico models. The intrinsically greater uncertainty of predictions for more complex apical endpoints from in vivo studies, such as acute and repeat-dose toxicity, is more challenging due to the number and diversity of mechanisms/ pathways potentially involved. To some extent, this hinders progress toward their acceptance as an alternative to animal testing. However, even for complex endpoints, the reliability of (Q)SAR-based predictions increases when the molecular target and mechanism of toxicity are known [107], suggesting that by considering a mechanistic understanding of a potential toxic effect in an assessment, uncertainty may be reduced and confidence in predictions based on in silico results will increase [108,109].

Increased understanding of pathways and key triggering mechanisms involved in the more complex types of toxicity, together with increased availability of in vitro data afford the opportunity to make a shift away from assessments based solely on descriptive endpoints, like the LD50, to mechanism-based endpoints that can be accurately assessed in vitro or by using in silico prediction models. The US NRC report "Toxicity Testing in the 21st Century: A Vision and a Strategy" [110] outlines the "new paradigm" based on the in vitro bioactivity assays using robotic high-throughput screening and supported by computational systems for modeling in vivo pharmacokinetics and distribution. Combining exposure modeling with toxicologic mechanism-based prediction models and an expert review creates a powerful paradigm with which to perform toxicity testing using an integrated approach to testing and assessment (IATA) [111]. Integration of existing in silico approaches and high throughput in vitro structure-based bioactivity information as part of a tiered testing strategy is helpful to reduce or eliminate animal testing for acute oral toxicity [112].

In addition to IST protocols for genetic toxicity [10] and for skin sensitization [11], discussion/position papers have been developed for neurotoxicity [14], carcinogenicity [16] and several primary organ systems (liver, kidney, lung and heart) [12,13]. As several of the topics are highly relevant to acute toxicity, those IST papers will provide more detailed information on in vitro endpoints and corresponding in silico models used to assess those commonly assessed target organs of toxicity.

2.4. In silico model considerations and assessment

In the previous section, several considerations were discussed within the context of computational models built from in vivo and in vitro data. The following summarizes these elements and includes some general considerations.

As part of the in silico toxicology protocol framework, Myatt and coworkers defined a series of considerations when electing to run an in silico model [9]. These include (1) the relevance of the predicted endpoint, (2) the validity and performance of the model, (3) the appropriateness of the chemical space the model has been trained on and is intended to support (i.e., the applicability domain), (4) whether the model can be combined with other models or information to increase reliability in the assessment, and (5) whether it can meaningfully support an expert review [9]. Each of these addresses a challenge recognized to potentially limit the usefulness of a model for an intended purpose. For example, models built using test/validation sets having a limited applicability domain may provide accurate predictions within narrow, localized areas of chemistry, but with decreased reliability outside the applicability domain.

The following additional aspects should be considered in the context of the prediction of acute oral toxicity for classification and labeling: i) whether the model's past performance has demonstrated consistent and accurate predictions of GHS classification for similar compounds to within the tolerance limit (e.g., within 1 category) needed and whether the model's tendency is to over- or underpredict; ii) the experiences others have had in applying it to similar scenarios; iii) evaluation of risk/benefit and development of a rationale for deciding whether the model is fit-for-purpose or can be utilized as a source of data in a weight-of-evidence approach. For example, if a model predicts a compound to be significantly toxic (e.g., GHS categories 1–3), how much confidence should be placed in the prediction? Conversely, if the model predicts a compound to be relatively non-toxic (e.g., category 5 or not classified), how much confidence should be given to that prediction? Likewise, it is important to put some thought into understanding why a particular prediction was made and to not simply take the prediction at face value.

One of the best ways of doing this is through an expert review of all the available information. If the compound is a pharmaceutical, for instance, and it has a mechanism of action known to be associated with highly toxic compounds (e.g., a microtubule stabilizer such as paclitaxel), it is important to acknowledge this information and to include it in a weight-of-evidence approach to support or refute the in silico prediction. To support such an expert review, the model should ideally be transparent and interpretable, providing associations between experiment-based acute toxicity data and the chemical class/characteristics being predicted. The model should provide a level of confidence for a prediction (e.g., in the form of a probability score for the prediction), and a way to assess and put into context the rationale or reasonableness of the prediction (e.g., by showing significant structural features and parameters the algorithm used to formulate a prediction). Ultimately, it is the availability of in vivo acute toxicity data that is one of the most important factors in generating high quality models to support classification and labelling. Finally, it should be considered whether models that provide some indication of potential toxicity at the whole organism level, such as the prediction of an LD50 value or GHS category, may be more fit-for-purpose in certain settings than models based entirely on in vitro data.

3. Use cases

How in silico predictions of acute oral toxicity can assist in different scenarios is demonstrated by different use cases listed in Table 7. Evaluation of data to support these use cases have been conducted [21,27]. For an example of assessment of acute toxicity using in vitro and New Approach Methods (NAMs) in rapid response situations, see Mumtaz et al., 2022 [24].

Table 7. Scenarios for the application of *in silico* predictions for acute oral toxicity.

Scenario	Description
Dose selection for *in vivo* studies	Utilizing *in silico* acute oral toxicity predictions to determine the study starting dose for an *in vivo* study.
GHS classification for safety data sheets	Utilizing *in silico* AOT predictions to determine the GHS classification for an unstudied compound.
Identify compounds which are dangerous goods	Utilizing *in silico* AOT predictions to inform whether a compound is dangerous good and, if so, what packing group it falls into.

3.1. Dose selection for nonclinical in vivo studies

Example scenarios: In vivo studies are required on each of the following compounds:

• Compound W has a predicted LD50 > 2000 mg/kg.

• Compound X has a predicted LD50 ≤ 5 mg/kg.

When a compound has limited data, initial dose selection for nonclinical studies can be a challenge. This applies to general in vivo toxicity studies as well as studies specifically geared toward elucidating acute toxicity. In silico models for the prediction of AOT can assist in determining the starting dose when in vivo studies are required or necessary. In the case of OECD-compliant AOT studies, the highest dose recommended is 2000 mg/kg [3,45–47]. When there is a lack of relevant data (as is the case for compounds generally being assessed in AOT studies), the recommended starting dose is 300 mg/kg.

In the case of Compound W with a predicted LD50 > 2000 mg/kg (GHS AOT category 5 or not classified), rather than starting the study at 300 mg/kg and subsequently dosing at 2000 mg/kg, the study can be conducted at the limit dose of 2000 mg/kg. If lethality is not observed, then the study is over and the minimum number of animals was utilized (i.e., dosing did not begin at 300 mg/kg and then move up to 2000 mg/ kg). If lethality and/or evident signs of toxicity are observed in the sighting study (one animal) then the dose can be decreased to 300 mg/ kg as is recommended and so on.

Similarly, if in the case of Compound X, the LD50 is predicted to be ≤ 5 mg/kg and therefore the starting dose of 5 mg/kg can be utilized and the subsequent dose determined based on any lethality and/or evident signs of toxicity at this dose. This approach can minimize unnecessary animal pain and suffering, which may have been observed if the dose levels were begun at the dose of 300 mg/kg and de-escalated accordingly (to 50 mg/kg and then to 5 mg/kg) if/when mortality is observed.

3.2. GHS classification in the context of hazard communication

Example scenario: Compound Y is being produced in the manufacturing facility and a safety data sheet is being prepared. There are questions regarding the acute toxicity hazard of the compound. How should the material be classified according to UN GHS (Table 2)?

When there is a lack of data regarding occupational hazards, advising on handling recommendations and personal protective equipment can prove difficult. Employees can be too protected (excess cost with no benefit) or inadequately protected (at high risk for adverse health effects from exposure). In silico approaches can inform hazard potential rapidly and predict the GHS AOT category that the compound may fall into. The in silico approaches may also inform the user of the analogues supporting the prediction, providing another means for data gathering. The GHS AOT category prediction can then be utilized to inform hazard potential, selection of appropriate personal protective equipment (PPE), and safe handling practices.

3.3. Identification of dangerous goods and packing group assignments

Example scenario: Compound Z needs to be shipped, and no data on its hazards are available to inform whether it is a dangerous good and, if so, what packing group it falls into.

A dangerous good is defined by the Federal Aviation Administration as a substance that is capable of posing an unreasonable risk to health, safety, and property when transported in commerce [113]. Compounds with an LD50 ≤ 300 mg/kg are considered dangerous goods (DGs) and the packaging group (PG) is assigned in accordance with Table 8. The identification of DGs is seen as the first step to reduce the risks posed by the product by defining the proper packaging, handling, and stowage.

Table 7. Correspondence between UN GHS categories and DOT packing groups

Rat oral LD50	UN GHS AOT Category[a]	DOT DG Packing Group[b]
≤ 5 mg/kg	1	I
5 < - ≤ 50 mg/kg	2	II
50 < - ≤ 300 mg/kg	3	III

[a] [3]
[b] [114]

In the absence of relevant data, the in silico AOT prediction can inform whether the compound may be a DG as well as falling within the recommended PG. For example, if Compound Z is predicted to fall within GHS Category 1, it can be interpreted as predicted to fall in PG I. Similarly, if the compound is predicted to fall within GHS Categories 4–5 or is not classified, and there is no available data to warrant consideration of the compound as a DG, then one could reasonably assume the material does not need to be classified as a DG.

3.4. Chemical hazard assessment of extractables or leachables

Example scenario: New container closure systems for pharmaceutical applications need to be assessed using extractable and leachable studies. Such studies are employed to evaluate container/closure systems to identify compounds that may contaminate API (Active Pharmaceutical Ingredient) over the course of the shelf life of a product. For many of the compounds detected, toxicology data is scant or unavailable. In such cases, a prediction of acute oral toxicity may be beneficial for determining whether a compound should be assigned a limit or beevaluated in confirmatory leachable studies. For example, irgaphos detergents can undergo chemical modifications resulting in degradants such as irgafos oxide or irgafos mono esters. An in silico evaluation of these molecules can give an indication of the potential for acute toxicity relative to the amounts observed in an extractable study. Upon evaluation, irgafos degradant compounds break out into GHS Cat II and IV (LD50 values >5–50 mg/kg and >300–2000 mg/kg) respectively. One approach might be that compounds that fall in Cat I or II (such as with irgafos oxide) could prompt additional evaluation. A comparison of the LD50 data to the predicted category for the irgafos compounds indicated that the model is conservative with respect to the in silico calls, which is appropriate when assessing chemical safety. While a parenteral acute toxicity model might be more relevant for extractable assessments, it is reasonable to evaluate compounds using the oral acute toxicity models, where the results could be used to inform risk assessment and the potential need for testing.

4. Conclusion

In silico methods are increasingly being used within chemical manufacturing- based industries to support assessments of toxicity from early discovery to product approval/registration, process manufacturing, and labeling. While in vivo animal studies remain the primary test system for meeting certain regulatory objectives for safety testing of compounds with a high risk of significant acute exposure, alternative testing and non-testing strategies based in vitro or in silico methods are increasingly used in certain other arenas. In silico statistical and expert knowledge-based models are widely used in weight-of-evidence assessment scenarios where direct testing in animals is not specifically indicated (such as for GHS classification and labeling). Moreover, as progress continues toward identifying additional toxicologic mechanisms and assay systems, opportunities to expand the coverage of in silico- based modeling endpoints will increase.

A significant obstacle to the wider acceptance of in silico-based results has been the general lack of guidance and standardization of minimum requirements needed to demonstrate performance, including assessing relevance (usefulness of the model for predicting the toxicological endpoint of interest), reliability (reflecting the quality of the information used for the assessment), and confidence in the ability of the approach to accurately predict the primary (i.e., apical) endpoint of interest, such as acute human toxicity [8]. Likewise, considerations for how to best document and report not only the results of an in silico assessment, but also the metadata used to describe processes related to developing, testing, identifying, managing, controlling, and applying systems have not been readily available.

The IST Protocol Consortium described by Myatt et al. [8,9] seeks to remedy the shortcomings of current guidances, which it is believed will foster increased understanding, acceptance, and reliance on in silico approaches for assessing toxicity. This paper achieves this objective by providing a framework for acute toxicity hazard assessment and by identifying discussion topics to consider when planning and reporting results derived from in silico technologies used to support a weight-ofevidence approach that may also include available experimental data and expert review.

5. Disclaimer

EPA Disclaimer

The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

CDC/ATSDR Disclaimer

The findings and conclusions in this article are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry.

B.A.T. Disclaimer

The authors declare that they have no affiliations with or involvement in any organization or individual with financial or non-financial interest in the subject matter or materials discussed in this manuscript. Ioanna Vardakou is employed by British American Tobacco.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:
Dr. Myatt reports grants from NIH during the conduct of the study. The remaining authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References
[1] J. Hamm, K. Sullivan, A.J. Clippinger, J. Strickland, S. Bell, B. Bhhatarai, B. Blaauboer, W. Casey, D. Dorman, A. Forsby, N. Garcia-Reyero, S. Gehen, R. Graepel, J. Hotchkiss, A. Lowit, J. Matheson, E. Reaves, L. Scarano, C.

Sprankle, J. Tunkel, D. Wilson, M. Xia, H. Zhu, D. Allen, Alternative approaches for identifying acute systemic toxicity: moving from research to regulatory testing, Toxicol. in Vitro 41 (2017) 245–259, https://doi.org/10.1016/j.tiv.2017.01.004.

[2] L.M. Aleksunes, D.L. Eaton, Principles of Toxicology, Chapter 2, in: C.D. Klaassen (Ed.), Casarett and Doull's Toxicology: The Basic Science of Poisons, 9th ed., McGraw-Hill Education, New York, NY, 2018. https://www.mheducation.com/highered/product/casarett-doull-s-toxicology-basicscience-poisons-9th-edition-klaassen/9781259863745.html.

[3] UN, Globally Harmonized System of Classification and Labelling of Chemicals (GHS), Ninth, United Nations, New York and Geneva, 2021. https://unece.org/transport/standards/transport/dangerous-goods/ghs-rev9-2021.

[4] J.W. Trevan, The error of determination of toxicity, Proc. R. Soc. B. 101 (1927) 483–514, https://doi.org/10.1098/rspb.1927.0030.

[5] E.O. Erhirhie, C.P. Ihekwereme, E.E. Ilodigwe, Advances in acute toxicity testing: strengths, weaknesses and regulatory acceptance, Interdiscip. Toxicol. 11 (2018) 5–12, https://doi.org/10.2478/intox-2018-0001.

[6] N. Sass, Humane endpoints and acute toxicity testing, ILAR J. 41 (2000) 114–123, https://doi.org/10.1093/ilar.41.2.114.

[7] D.P. Russo, J. Strickland, A.L. Karmaus, W. Wang, S. Shende, T. Hartung, L.M. Aleksunes, H. Zhu, Nonanimal models for acute toxicity evaluations: Applying data-driven profiling and read-across, Environ. Health Perspect. 127 (2019) 047001. https://doi.org/10.1289/EHP3614.

[8] G.J. Myatt, A. Bassan, D. Bower, K.M. Crofton, K.P. Cross, J.C. Graham, C. Hasselgren, R.A. Jolly, S. Miller, M. Pavan, R.R. Tice, C. Zwickl, C. Johnson, Increasing the acceptance of in silico toxicology through development of protocols and position papers, Comput. Toxicol. 21 (2022) 100209, https://doi.org/10.1016/j.comtox.2021.100209.

[9] G.J. Myatt, E. Ahlberg, Y. Akahori, D. Allen, A. Amberg, L.T. Anger, A. Aptula, S. Auerbach, L. Beilke, P. Bellion, R. Benigni, J. Bercu, E.D. Booth, D. Bower, A. Brigo, N. Burden, Z. Cammerer, M.T.D. Cronin, K.P. Cross, L. Custer, M. Dettwiler, K. Dobo, K.A. Ford, M.C. Fortin, S.E. Gad-McDonald, N. Gellatly, V. Gervais, K.P. Glover, S. Glowienke, J. Van Gompel, S. Gutsell, B. Hardy, J.S. Harvey, J. Hillegass, M. Honma, J.-H. Hsieh, C.-W. Hsu, K. Hughes, C. Johnson, R. Jolly, D. Jones, R. Kemper, M.O. Kenyon, M.T. Kim, N.L. Kruhlak, S.A. Kulkarni, K. Kümmerer, P. Leavitt, B. Majer, S. Masten, S. Miller, J. Moser, M. Mumtaz, W. Muster, L. Neilson, T.I. Oprea, G. Patlewicz, A. Paulino, E. Lo Piparo, M. Powley, D.P. Quigley, M.V. Reddy, A.-N. Richarz, P. Ruiz, B. Schilter, R. Serafimova, W. Simpson, L. Stavitskaya, R. Stidl, D. Suarez-Rodriguez, D.T. Szabo, A. Teasdale, A. Trejo-Martin,

J.-P. Valentin, A. Vuorinen, B.A. Wall, P. Watts, A.T. White, J. Wichard, K.L. Witt, A. Woolley, D. Woolley, C. Zwickl, C. Hasselgren, In silico toxicology protocols, Regul. Toxicol. Pharmacol. 96 (2018) 1–17, https://doi.org/10.1016/j.yrtph.2018.04.014.

[10] C. Hasselgren, E. Ahlberg, Y. Akahori, A. Amberg, L.T. Anger, F. Atienzar, S. Auerbach, L. Beilke, P. Bellion, R. Benigni, J. Bercu, E.D. Booth, D. Bower, A. Brigo, Z. Cammerer, M.T.D. Cronin, I. Crooks, K.P. Cross, L. Custer, K. Dobo, T. Doktorova, D. Faulkner, K.A. Ford, M.C. Fortin, M. Frericks, S.E. Gad-McDonald, N. Gellatly, H. Gerets, V. Gervais, S. Glowienke, J. Van Gompel, J.S. Harvey, J. Hillegass, M. Honma, J.-H. Hsieh, C.-W. Hsu, T.S. Barton-Maclaren, C. Johnson, R. Jolly, D. Jones, R. Kemper, M.O. Kenyon, N.L. Kruhlak, S.A. Kulkarni, K. Kümmerer, P. Leavitt, S. Masten, S. Miller, C. Moudgal, W. Muster, A. Paulino, E. Lo Piparo, M. Powley, D.P. Quigley, M.V. Reddy, A.-N. Richarz, B. Schilter, R.D. Snyder, L. Stavitskaya, R. Stidl, D.T. Szabo, A. Teasdale, R.R. Tice, A. Trejo - Martin, A. Vuorinen, B.A. Wall, P. Watts, A.T. White, J. Wichard, K.L. Witt, A. Woolley, D. Woolley, C. Zwickl, G.J. Myatt, Genetic toxicology in silico protocol, Regul. Toxicol. Pharmacol. 107 (2019) 104403, https://doi.org/10.1016/j.yrtph.2019.104403.

[11] C. Johnson, E. Ahlberg, L.T. Anger, L. Beilke, R. Benigni, J. Bercu, S. Bobst, D. Bower, A. Brigo, S. Campbell, M.T.D. Cronin, I. Crooks, K.P. Cross, T. Doktorova, T. Exner, D. Faulkner, I.M. Fearon, M. Fehr, S.C. Gad, V. Gervais, A. Giddings, S. Glowienke, B. Hardy, C. Hasselgren, J. Hillegass, R. Jolly, E. Krupp, L. Lomnitski, J. Magby, J. Mestres, L. Milchak, S. Miller, W. Muster, L. Neilson, R. Parakhia, A. Parenty, P. Parris, A. Paulino, A.T. Paulino, D.W. Roberts, H. Schlecker, R. Stidl, D. Suarez-Rodrigez, D.T. Szabo, R.R. Tice, D. Urbisch, A. Vuorinen, B. Wall, T. Weiler, A.T. White, J. Whritenour, J. Wichard, D. Woolley, C. Zwickl, G.J. Myatt, Skin sensitization in silico protocol, Regul. Toxicol. Pharmacol. 116 (2020) 104688, https://doi.org/10.1016/j.yrtph.2020.104688.

[12] A. Bassan, V.M. Alves, A. Amberg, L.T. Anger, L. Beilke, A. Bender, A. Bernal, M.T.D. Cronin, J.-H. Hsieh, C. Johnson, R. Kemper, M. Mumtaz, L. Neilson, M. Pavan, A. Pointon, J. Pletz, P. Ruiz, D.P. Russo, Y. Sabnis, R. Sandhu, M. Schaefer, L. Stavitskaya, D.T. Szabo, J.-P. Valentin, D. Woolley, C. Zwickl, G.J. Myatt, In silico approaches in organ toxicity hazard assessment: current status and future needs for predicting heart, kidney and lung toxicities, Comput. Toxicol. 20 (2021) 100188.

[13] A. Bassan, V.M. Alves, A. Amberg, L.T. Anger, S. Auerbach, L. Beilke, A. Bender, M.T.D. Cronin, K.P. Cross, J.-H. Hsieh, N. Greene, R. Kemper, M.T. Kim, M. Mumtaz, T. Noeske, M. Pavan, J. Pletz, D.P. Russo, Y. Sabnis, M. Schaefer, D.T. Szabo, J.-P. Valentin, J. Wichard, D. Williams, D. Woolley, C. Zwickl, G.J. Myatt, In silico approaches in organ toxicity hazard assessment: current status and future needs in predicting liver toxicity, Comput. Toxicol. 20 (2021) 100187.

[14] K. Crofton, A. Bassan, M. Behl, Y. Chushak, E. Fritsche, J. Gearhart, S. Marty, M. Mumtaz, M. Pavan, P. Ruiz, T. Shaffer, M. Sachana, R. Selvam, L. Stavitskaya, D. Szabo, R. Tice, D. Wilson, D. Woolley, G.J. Myatt, Current status and future needs for a neurotoxicity hazard assessment framework that integrates in silico approaches, (2022) Submitted.

[15] C. Johnson, L.T. Anger, R. Benigni, D. Bower, F. Bringezu, K.M. Crofton, M.T.D. Cronin, K.P. Cross, M. Dettwiler, M. Frericks, F. Melnikov, S. Miller, D.W. Roberts, D. Suarez-Rodrigez, A. Roncaglioni, E. Lo Piparo, R.R. Tice, C. Zwickl, G.J. Myatt, Evaluating confidence in toxicity assessments based on experimental data and in silico predictions, Comput. Toxicol. 21 (2022) 100204, https://doi.org/10.1016/j.comtox.2021.100204.

[16] R.R. Tice, A. Bassan, A. Amberg, L.T. Anger, M.A. Beal, P. Bellion, R. Benigni, J. Birmingham, A. Brigo, F. Bringezu, L. Ceriani, I. Crooks, K. Cross, R. Elespuru, D. Faulkner, M.C. Fortin, P. Fowler, M. Frericks, H.H.J. Gerets, G.D. Jahnke, D.R. Jones, N.L. Kruhlak, E. Lo Piparo, J. Lopez-Belmonte, A. Luniwal, A. Luu, F. Madia, S. Manganelli, B. Manickam, J. Mestres, A.L. Mihalchik-Burhans, L. Neilson, A. Pandiri, M. Pavan, C.V. Rider, J.P. Rooney, A. Trejo-Martin, K.H. Watanabe-Sailor, A.T. White, D. Woolley, G.J. Myatt, In silico approaches in carcinogenicity hazard assessment: current status and future needs, Comput. Toxicol. 20 (2021) 100191, https://doi.org/10.1016/j.comtox.2021.100191.

[17] S. Ali, H.G.J.V. Mil, M.K. Richardson, F. Mueller, Large-scale assessment of the zebrafish embryo as a possible predictive model in toxicity testing, PLoS One. 6 (6) (2011) e21076, https://doi.org/10.1371/journal.pone.0021076.

[18] N.A. Ducharme, D.M. Reif, J.-A. Gustafsson, M. Bondesson, Comparison of toxicity values across zebrafish early life stages and mammalian studies: implications for chemical testing, Reprod. Toxicol. 55 (2015) 3–10, https://doi.org/10.1016/j.reprotox.2014.09.005.

[19] P.R. Hunt, The C. elegans model in toxicity testing, J. Appl. Toxicol. 37 (2017) 50–59, https://doi.org/10.1002/jat.3357.

[20] NASEM, National Academies of Sciences, Engineering, and Medicine - Application of Modern Toxicology Approaches for Predicting Acute Toxicity for Chemical Defense, National Academies Press, Washington, D.C., 2015. https://doi.org/10.17226/21775.

[21] J. Bercu, M.J. Masuda-Herrera, A. Trejo-Martin, C. Hasselgren, J. Lord, J. Graham, M. Schmitz, L. Milchak, C. Owens, S.H. Lal, R.M. Robinson, S. Whalley, P. Bellion, A. Vuorinen, K. Gromek, W.A. Hawkins, I. van de Gevel, K. Vriens, R. Kemper, R. Naven, P. Ferrer, G.J. Myatt, A cross-industry collaboration to assess if acute oral toxicity (Q)SAR models are fit-for-purpose for GHS classification and labelling, Regul. Toxicol. Pharmacol. 120 (2021) 104843, https://doi.org/10.1016/j.yrtph.2020.104843.

[22] S. Creton, I.C. Dewhurst, L.K. Earl, S.C. Gehen, R.L. Guest, J.A. Hotchkiss, I. Indans, M.R. Woolhiser, R. Billington, Acute toxicity testing of chemicals-Opportunities to avoid redundant testing and use alternative approaches, Crit. Rev. Toxicol. 40 (2010) 50–83, https://doi.org/10.3109/10408440903401511.

[23] A. Schrage, K. Hempel, M. Schulz, S.N. Kolle, B. van Ravenzwaay, R. Landsiedel, Refinement and reduction of acute oral toxicity testing: a critical review of the use of cytotoxicity data, Altern. Lab. Anim. 39 (2011) 273–295, https://doi.org/10.1177/026119291103900311.

[24] M.M. Mumtaz, R.A. Nickle, J.C. Lambert, M.S. Johnson, Advances in assessing hazard and risk to emerging threats and emergency response: comparing and contrasting efforts of 3 federal agencies, Toxicol. Sci. 185 (2022) 1–9, https://doi.org/10.1093/toxsci/kfab126.

[25] J. Strickland, A.J. Clippinger, J. Brown, D. Allen, A. Jacobs, J. Matheson, A. Lowit, E.N. Reinke, M.S. Johnson, M.J. Quinn, D. Mattie, S.C. Fitzpatrick, S. Ahir, N. Kleinstreuer, W. Casey, Status of acute systemic toxicity testing requirements and data uses by U.S. regulatory agencies, Regul. Toxicol. Pharmacol. 94 (2018) 183–196, https://doi.org/10.1016/j.yrtph.2018.01.022.

[26] ECHA, Guidance on information requirements and chemical safety assessment. Chapter R.6: QSARs and grouping of chemicals, 2008. https://echa.europa.eu/documents/10162/13632/information_requirements_r6_en.pdf.

[27] J.C. Graham, M. Rodas, J. Hillegass, G. Schulze, The performance, reliability and potential application of in silico models for predicting the acute oral toxicity of pharmaceutical compounds, Regul. Toxicol. Pharmacol. 119 (2021) 104816, https://doi.org/10.1016/j.yrtph.2020.104816.

[28] ECHA, Advice on skin and eye irritation testing helps reduce animal tests, All news - ECHA. (2016). https://echa.europa.eu/it/-/advice-on-skin-and-eyeirritation-testing-helps-reduce-animal-tests (accessed February 1, 2022).

[29] ECHA, Guidance on information requirements and chemical safety assessment Chapter R.7a: endpoint specific guidance. Version 6.0, Publications Office of the EU, 2017. https://doi.org/10.2823/337352.

[30] JRC, EURL ECVAM dataset on alternative methods to animal experimentation (DB-ALM). European Commission, Joint Research Centre (JRC), (2019). http://data.europa.eu/89h/b7597ada-148d-4560-9079-ab0a5539cad3.

[31] W.S. Stokes, S. Casati, J. Strickland, M. Paris, Neutral red uptake cytotoxicity tests for estimating starting doses for acute oral toxicity tests, Curr. Protocols Toxicol. 36 (2008) 20.4.1-20.4.20. https://doi.org/10.1002/0471140856.tx2004s36.

[32] R. Huang, M. Xia, S. Sakamuru, J. Zhao, S.A. Shahane, M. Attene-Ramos, T. Zhao, C.P. Austin, A. Simeonov, Modelling the Tox21 10K chemical profiles for in vivo toxicity prediction and mechanism characterization, Nat. Commun. 7 (2016) 10425, https://doi.org/10.1038/ncomms10425.

[33] C.S. Pridgeon, C. Schlott, M.W. Wong, M.B. Heringa, T. Heckel, J. Leedale, L. Launay, V. Gryshkova, S. Przyborski, R.N. Bearon, E.L. Wilkinson, T. Ansari, J. Greenman, D.F.G. Hendriks, S. Gibbs, J. Sidaway, R.L. Sison-Young, P. Walker, M.J. Cross, B.K. Park, C.E.P. Goldring, Innovative organotypic in vitro models for safety assessment: aligning with regulatory requirements and understanding models of the heart, skin, and liver as paradigms, Arch. Toxicol. 92 (2018) 557–569, https://doi.org/10.1007/s00204-018-2152-9.

[34] P. Prieto, R. Graepel, K. Gerloff, L. Lamon, M. Sachana, F. Pistollato, L. Gribaldo, A. Bal-Price, A. Worth, Investigating cell type specific mechanisms contributing to acute oral toxicity, ALTEX 36 (2019) 39–64. https://doi.org/10.14573/altex.1805181.

[35] N.S. Sipes, M.T. Martin, P. Kothiya, D.M. Reif, R.S. Judson, A.M. Richard, K.A. Houck, D.J. Dix, R.J. Kavlock, T.B. Knudsen, Profiling 976 ToxCast chemicals across 331 enzymatic and receptor signaling assays, Chem. Res. Toxicol. 26 (2013) 878–895, https://doi.org/10.1021/tx400021f.

[36] M.-G. Zurich, S. Stanzel, A. Kopp-Schneider, P. Prieto, P. Honegger, Evaluation of aggregating brain cell cultures for the detection of acute organ-specific toxicity, Toxicol. In Vitro. 27 (2013) 1416–1424, https://doi.org/10.1016/j.tiv.2012.06.018.

[37] UNCED, United Nations Conference on Environment and Development - Environmentally sound management of toxic chemicals, including prevention of illegal international traffic in toxic and dangerous products. Agenda 21, Chapter 19., (1992). https://www.ilo.org/legacy/english/protection/safework/ghs/ghsdocs/chapt19.pdf.

[38] UN, Globally Harmonized System of Classification and Labelling of Chemicals (GHS), First revised edition, United Nations, New York and Geneva, 2005. https://unece.org/ghs-rev1-2005.

[39] UNITAR, Understanding the Globally Harmonized System of Classification and Labelling of Chemicals (GHS). A companion guide to the GHS Purple Book., (2012). https://cwm.unitar.org/national-profiles/publications/cw/ghs/GHS_Companion_Guide_final_June2012_EN.pdf.

[40] OSHA, Hazard Communication, United States Department of Labor. (2022). https://www.osha.gov/hazcom (accessed February 1, 2022).

[41] J.W. Firman, M.T.D. Cronin, P.H. Rowe, E. Semenova, J.E. Doe, The use of Bayesian methodology in the development and validation of a tiered assessment approach towards prediction of rat acute oral toxicity, Arch. Toxicol. 96 (2022) 817–830, https://doi.org/10.1007/s00204-021-03205-x.

[42] H.-J. Klimisch, M. Andreae, U. Tillmann, A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data, Regul. Toxicol. Pharmacol. 25 (1997) 1–5, https://doi.org/10.1006/rtph.1996.1076.

[43] S. Kutsarova, T.W. Schultz, A. Chapkanov, D. Cherkezova, A. Mehmed, S. Stoeva, C. Kuseva, D. Yordanova, M. Georgiev, T. Petkov, O.G. Mekenyan, The QSAR Toolbox automated read-across workflow for predicting acute oral toxicity: II. Verification and validation, Comput. Toxicol. 20 (2021) 100194, https://doi.org/10.1016/j.comtox.2021.100194.

[44] OECD, Guidance document on acute oral toxicity testing, OECD Publishing, Paris, 2001. https://doi.org/10.1787/9789264078413-en.

[45] OECD, Test No. 425: Acute Oral Toxicity: Up-and-Down Procedure, OECD Publishing, Paris, 2008. https://doi.org/10.1787/9789264071049-en.

[46] OECD, Test No. 423: Acute Oral toxicity - Acute Toxic Class Method, OECD Publishing, Paris, 2002. https://doi.org/10.1787/9789264071001-en.

[47] OECD, Test No. 420: Acute Oral Toxicity - Fixed Dose Procedure, OECD Publishing, Paris, 2002. https://doi.org/10.1787/9789264070943-en.

[48] OECD, Test No. 401: Acute Oral Toxicity. (Following the OECD Council decision, the test 401 'Acute Oral Toxicity' was deleted on 17th December 2002), OECD Publishing, Paris, 1987. https://doi.org/10.1787/9789264040113-en.

[49] OECD, Test No. 402: Acute Dermal Toxicity, OECD Publishing, Paris, 2017. https://doi.org/10.1787/9789264070585-en.

[50] OECD, Test No. 403: Acute Inhalation Toxicity, OECD Publishing, Paris, 2009. https://doi.org/10.1787/9789264070608-en.

[51] OECD, Test No. 436: Acute Inhalation Toxicity – Acute Toxic Class Method, OECD Publishing, Paris, 2009. https://doi.org/10.1787/9789264076037-en.

[52] OECD, Test No. 433: Acute Inhalation Toxicity: Fixed Concentration Procedure, OECD Publishing, Paris, 2018. https://doi.org/10.1787/9789264284166-en.

[53] R. Bureau, Nontest Methods to Predict Acute Toxicity: State of the Art for Applications of In Silico Methods, in: O. Nicolotti (Ed.), Computational Toxicology, First, Humana Press, New York, NY, 2018: pp. 519–534. https://doi.org/10.1007/978-1-4939-7899-1_24.

[54] J. Burton, A.P. Worth, I. Tsakovska, A. Diukendjieva, In silico models for acute systemic toxicity, in: E. Benfenati (Ed.), In Silico Methods for Predicting Drug Toxicity, Humana Press, New York, NY, 2016: pp. 177–200. https://doi.org/10.1007/978-1-4939-3609-0_10.

[55] M.T.D. Cronin, J.C. Dearden, QSAR in toxicology. 2. Prediction of acute mammalian toxicity and interspecies correlations, Quant. Struct.-Act. Relat. 14 (2) (1995) 117–120.

[56] R. Gonella Diaza, S. Manganelli, A. Esposito, A. Roncaglioni, A. Manganaro, E. Benfenati, Comparison of in silico tools for evaluating rat oral acute toxicity, SAR QSAR Environ. Res. 26 (2015) 1–27, https://doi.org/10.1080/1062936X.2014.977819.

[57] I. Tsakovska, A. Diukendjieva, A.P. Worth, In Silico Models for Predicting Acute Systemic Toxicity, in: E. Benfenati (Ed.), In Silico Methods for Predicting Drug Toxicity, Humana, New York, NY, 2022: pp. 259–289. https://doi.org/10.1007/ 978-1-0716-1960-5_12.

[58] I. Tsakovska, I. Lessigiarska, T. Netzeva, A.P. Worth, A mini review of mammalian toxicity (Q)SAR models, QSAR Comb. Sci. 27 (2008) 41–48, https://doi.org/10.1002/qsar.200710107.

[59] I. Tsakovska, I. Lessigiarska, T. Netzeva, A. Worth, Review of (Q)SARs for mammalian toxicity, Bioautomation. 5 (2006) 90–105. https:// publications.jrc.ec.europa.eu/repository/handle/JRC35954.

[60] R.D. Benz, H. Irausquin, Priority-based assessment of food additives database of the U.S. Food and Drug Administration Center for Food Safety and Applied Nutrition, Environ. Health Perspect. 96 (1991) 85–89, https://doi.org/10.1289/ehp.919685.

[61] A.L. Karmaus, Rat Oral Acute Toxicity Database and Evaluation of Variability, presented at the NTP Workshop on Predictive Models for Acute Oral Systemic Toxicity, April 11 2018, (2018). https://ntp.niehs.nih.gov/iccvam/meetings/atmodels-2018/ppt/4-karmaus.pdf.

[62] K. Mansouri, A.L. Karmaus, J. Fitzpatrick, G. Patlewicz, P. Pradeep, D. Alberga, N. Alepee, T.E.H. Allen, D. Allen, V.M. Alves, C.H. Andrade, T.R. Auernhammer, D. Ballabio, S. Bell, E. Benfenati, S. Bhattacharya, J.V. Bastos, S. Boyd, J.B. Brown, S.J. Capuzzi, Y. Chushak, H. Ciallella, A.M. Clark, V. Consonni, P.R. Daga, S. Ekins, S. Farag, M. Fedorov, D. Fourches, D. Gadaleta, F. Gao, J.M. Gearhart, G. Goh, J.M. Goodman, F. Grisoni, C.M. Grulke, T. Hartung, M. Hirn, P. Karpov, A. Korotcov, G.J. Lavado, M. Lawless, X. Li, T. Luechtefeld, F. Lunghini, G.F. Mangiatordi, G. Marcou, D. Marsh, T. Martin, A. Mauri, E.N. Muratov, G.J. Myatt, D.-T. Nguyen, O. Nicolotti, R. Note, P. Pande, A.K. Parks, T. Peryea, A.H. Polash, R. Rallo, A. Roncaglioni, C. Rowlands, P. Ruiz, D.P. Russo, A. Sayed, R. Sayre, T. Sheils, C. Siegel, A.C. Silva, A. Simeonov, S. Sosnin, N. Southall, J. Strickland, Y. Tang, B. Teppen, I.V. Tetko, D. Thomas, V. Tkachenko, R. Todeschini, C. Toma, I. Tripodi, D. Trisciuzzi, A. Tropsha, A. Varnek, K. Vukovic, Z. Wang, L. Wang, K.M. Waters, A.J. Wedlake, S.J. Wijeyesakere, D. Wilson, Z. Xiao, H. Yang, G.

Zahoranszky-Kohalmi, A.V. Zakharov, F.F. Zhang, Z. Zhang, T. Zhao, H. Zhu, K.M. Zorn, W. Casey, N.C. Kleinstreuer, CATMoS: collaborative acute toxicity modelling suite, Environ Health Perspect. 129 (2021) 47013, https://doi.org/10.1289/EHP8495.

[63] NIOSH, Registry of Toxic Effects of Chemical Substances (RTECS). Comprehensive guide to the RTECS., U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, 1997. https://doi.org/10.26616/NIOSHPUB97119.

[64] N.C. Kleinstreuer, A.L. Karmaus, K. Mansouri, D.G. Allen, J.M. Fitzpatrick, G. Patlewicz, Predictive models for acute oral systemic toxicity: a workshop to bridge the gap from research to regulation, Comput. Toxicol. 8 (2018) 21–24, https:// doi.org/10.1016/j.comtox.2018.08.002.

[65] T. Fan, G. Sun, L. Zhao, X. Cui, R. Zhong, QSAR and classification study on prediction of acute oral toxicity of N-Nitroso compounds, Int. J. Mol. Sci. 19 (2018) 3015, https://doi.org/10.3390/ijms19103015.

[66] P. Ruiz, G. Begluitti, T. Tincher, J. Wheeler, M. Mumtaz, Prediction of acute mammalian toxicity using QSAR methods: a case study of sulfur mustard and its breakdown products, Molecules 17 (2012) 8982–9001, https://doi.org/10.3390/molecules17088982.

[67] B. Rasulev, H. Kušić, D. Leszczynska, J. Leszczynski, N. Koprivanac, QSAR modeling of acute toxicity on mammals caused by aromatic compounds: the case study using oral LD50 for rats, J. Environ. Monit. 12 (2010) 1037, https://doi.org/10.1039/b919489d.

[68] R. García-Domenech, P. Alarcón-Elbal, G. Bolas, R. Bueno-Marí, F.A. Chordá-Olmos, S.A. Delacour, M.C. Mouriño, A. Vidal, J. Gálvez, Prediction of acute toxicity of organophosphorus pesticides using topological indices, SAR QSAR Environ. Res. 18 (2007) 745–755, https://doi.org/10.1080/10629360701698712.

[69] Y.K. Koleva, M.T.D. Cronin, J.C. Madden, J.A.H. Schwöbel, Modelling acute oral mammalian toxicity. 1. Definition of a quantifiable baseline effect, Toxicol. in Vitro. 25 (2011) 1281–1293, https://doi.org/10.1016/j.tiv.2011.04.015.

[70] S. Chavan, I. Nicholls, B. Karlsson, A. Rosengren, D. Ballabio, V. Consonni, R. Todeschini, Towards global QSAR model building for acute toxicity: munro database case study, Int. J. Mol. Sci. 15 (2014) 18162–18174, https://doi.org/10.3390/ijms151018162.

[71] H. Zhu, T.M. Martin, L. Ye, A. Sedykh, D.M. Young, A. Tropsha, Quantitative structure–activity relationship modeling of rat acute toxicity by oral exposure, Chem. Res. Toxicol. 22 (2009) 1913–1921, https://doi.org/10.1021/tx900189p.

[72] H. Zhu, L. Ye, A. Richard, A. Golbraikh, F.A. Wright, I. Rusyn, A. Tropsha, A novel two-step hierarchical quantitative structure–activity relationship modelling work flow for predicting acute toxicity of chemicals in rodents, Environ. Health Perspect. 117 (2009) 1257–1264, https://doi.org/10.1289/ehp.0800471.

[73] D. Gadaleta, K. Vuković, C. Toma, G.J. Lavado, A.L. Karmaus, K. Mansouri, N.C. Kleinstreuer, E. Benfenati, A. Roncaglioni, SAR and QSAR modeling of a large collection of LD50 rat acute oral toxicity data, J. Cheminform. 11 (2019) 58, https://doi.org/10.1186/s13321-019-0383-2.

[74] C.R. García-Jacas, Y. Marrero-Ponce, F. Cortés-Guzmán, J. Suárez-Lezcano, F.O. Martinez-Rios, L.A. García-González, M. Pupo-Meriño, K. Martinez-Mayorga, Enhancing acute oral toxicity predictions by using consensus modeling and algebraic form-based 0D-to-2D molecular encodes, Chem. Res. Toxicol. 32 (2019) 1178–1192, https://doi.org/10.1021/acs.chemrestox.9b00011.

[75] T. Lei, Y. Li, Y. Song, D. Li, H. Sun, T. Hou, ADMET evaluation in drug discovery: 15. Accurate prediction of rat oral acute toxicity using relevance vector machine and consensus modeling, J. Cheminf. 8 (2016) 6, https://doi.org/10.1186/s13321-016-0117-7.

[76] T. Luechtefeld, UL Cheminformatics Suite, presented at the NTP Workshop on Predictive Models for Acute Oral Systemic Toxicity, April 11 2018, (2018). https://ntp.niehs.nih.gov/iccvam/meetings/at-models-2018/ppt/9-luechtefeld.pdf.

[77] F. Lunghini, G. Marcou, P. Azam, D. Horvath, R. Patoux, E. Van Miert, A. Varnek, Consensus models to predict oral rat acute toxicity and validation on a dataset coming from the industrial context, SAR QSAR Environ. Res. 30 (2019) 879–897, https://doi.org/10.1080/1062936X.2019.1672089.

[78] A.A. Sayed, Acute oral systemic toxicity: Consensus approach for modelling acute systemic oral toxicity and LD50 data using machine learning and in silico approaches, presented at the NTP Workshop on Predictive Models for Acute Oral Systemic Toxicity, April 11 2018, (2018). https://ntp.niehs.nih.gov/iccvam/meetings/at-models-2018/ppt/14-sayed.pdf.

[79] V. Kleandrova, F. Luan, A. Speck-Planche, M. Cordeiro, In silico assessment of the acute toxicity of chemicals: recent advances and new model for multitasking prediction of toxic effect, Mini-Rev. Med. Chem. 15 (2015) 677–686, https://doi.org/10.2174/1389557515666150219143604.

[80] M. Lawless, P.R. Daga, M. Waldman, R. Fraczkiewicz, R.D. Clarke, J. DiBella, M.B. Bolger, Predicting Five Rat Acute Toxicity Endpoints with ANNE Models, presented at the NTP Workshop on Predictive Models for Acute Oral Systemic Toxicity, April 11 2018, (2018). https://ntp.niehs.nih.gov/iccvam/meetings/atmodels-2018/ppt/15-lawless.pdf.

[81] S. Jain, V.B. Siramshetty, V.M. Alves, E.N. Muratov, N. Kleinstreuer, A. Tropsha, M.C. Nicklaus, A. Simeonov, A.V. Zakharov, Large-scale modeling of multispecies acute toxicity end points using consensus of multitask deep learning methods, J.Chem. Inf. Model. 61 (2021) 653–663, https://doi.org/10.1021/acs.jcim.0c01164.

[82] R. Liu, M. Madore, K.P. Glover, M.G. Feasel, A. Wallqvist, Assessing deep and shallow learning methods for quantitative prediction of acute chemical toxicity, Toxicol. Sci. 164 (2018) 512–526, https://doi.org/10.1093/toxsci/kfy111.

[83] A. Zakharov, Multitask Deep learning modelling of rodent acute toxicity, presented at the NTP Workshop on Predictive Models for Acute Oral Systemic Toxicity, April 11 2018, (2018). https://ntp.niehs.nih.gov/iccvam/meetings/atmodels-2018/ppt/11-zakharov.pdf.

[84] J. Lu, J. Peng, J. Wang, Q. Shen, Y. Bi, L. Gong, M. Zheng, X. Luo, W. Zhu, H. Jiang, K. Chen, Estimation of acute oral toxicity in rat using local lazy learning, J. Cheminf. 6 (2014) 26, https://doi.org/10.1186/1758-2946-6-26.

[85] A. Roncaglioni, C. Toma, G. Lavado, K. Vuković, D. Gadaleta, E. Benfenati, Modeling quantitative acute oral systemic toxicity based on a k-Nearest Neighbor (k-NN) algorithm, presented at the NTP Workshop on Predictive Models for Acute Oral Systemic Toxicity, April 11 2018, (2018). https://ntp.niehs.nih.gov/iccvam/meetings/at-models-2018/ppt/7-roncaglioni.pdf.

[86] O.A. Raevsky, V.J. Grigor'ev, E.A. Modina, A.P. Worth, Prediction of acute toxicity to mice by the Arithmetic Mean Toxicity (AMT) modelling approach, SAR QSAR Environ. Res. 21 (3–4) (2010) 265–275.

[87] G.J. Myatt, Leadscope predictive models for acute oral systemic toxicity, presented at the NTP Workshop on Predictive Models for Acute Oral Systemic Toxicity, April 11 2018, (2018). https://ntp.niehs.nih.gov/iccvam/meetings/atmodels-2018/ppt/12-myatt.pdf.

[88] E. Muratov, V.M. Alves, S.J. Capuzzi, S. Farag, J. Borba, A. Carvalho, C.H. Andrade, A. Tropsha, NTP Collaboration on Modeling Acute Systemic Oral Toxicity: Results and Ideas, presented at the NTP Workshop on Predictive Models for Acute Oral Systemic Toxicity, April 11 2018, (2018). https://ntp.niehs.nih.gov/iccvam/meetings/at-models-2018/ppt/6-muratov.pdf.

[89] Z. Zhang, H. Wan, Y. Wang, H. Xin, P. Yu, Y. Li, S. Gehen, A clustering-based QSAR model for acute oral systemic toxicity, presented at the NTP Workshop on Predictive Models for Acute Oral Systemic Toxicity, April 11 2018, (2018). https://ntp.niehs.nih.gov/iccvam/meetings/at-models-2018/ppt/13-zhang.pdf.

[90] A. Sazonovas, P. Japertas, R. Didziapetris, Estimation of reliability of predictions and model applicability domain evaluation in the analysis of acute toxicity (LD50), SAR QSAR Environ. Res. 21 (2010) 127–148, https://doi.org/10.1080/10629360903568671.

[91] T. Luechtefeld, D. Marsh, C. Rowlands, T. Hartung, Machine learning of toxicological big data enables Read-Across Structure Activity Relationships (RASAR) outperforming animal test reproducibility, Toxicol. Sci. 165 (2018) 198–212, https://doi.org/10.1093/toxsci/kfy152.

[92] M. Hamadache, O. Benkortbi, S. Hanini, A. Amrane, L. Khaouane, C. Si Moussa, A quantitative structure activity relationship for acute oral toxicity of pesticides on rats: validation, domain of application and prediction, J. Hazard. Mater. 303 (2016) 28–40, https://doi.org/10.1016/j.jhazmat.2015.09.021.

[93] R. Liu, K.P. Glover, M.G. Feasel, A. Wallqvist, General approach to estimate error bars for quantitative structure–activity relationship predictions of molecular activity, J. Chem. Inf. Model. 58 (2018) 1561–1575, https://doi.org/10.1021/acs.jcim.8b00114.

[94] D. Ballabio, F. Grisoni, V. Consonni, R. Todeschini, Integrated QSAR models to predict acute oral systemic toxicity, Mol. Inf. 38 (2018) 1800124, https://doi.org/10.1002/minf.201800124.

[95] A. Lagunin, A. Zakharov, D. Filimonov, V. Poroikov, QSAR modelling of rat acute toxicity on the basis of PASS prediction, Mol. Inf. 30 (2011) 241–250, https://doi.org/10.1002/minf.201000151.

[96] K. Vukovic, D. Gadaleta, E. Benfenati, Methodology of aiQSAR: a group-specific approach to QSAR modelling, J. Cheminf. 11 (2019) 27, https://doi.org/10.1186/s13321-019-0350-y.

[97] NTP, Workshop: Predictive Models for Acute Oral Systemic Toxicity, William H. Natcher Conference Center, National Institutes of Health Bethesda, Maryland, USA, 2018. https://ntp.niehs.nih.gov/whatwestudy/niceatm/3rs-meetings/pastmeetings/tox-models-2018/index.html.

S. Hoffmann, A. Kinsner-Ovaskainen, P. Prieto, I. Mangelsdorf, C. Bieler, T. Cole, Acute oral toxicity: variability, reliability, relevance and interspecies comparison of rodent LD50 data from literature surveyed for the ACuteTox project, Regul. Toxicol. Pharmacol. 58 (2010) 395–407, https://doi.org/10.1016/j.yrtph.2010.08.004.

[99] L.A. Low, D.A. Tagle, Tissue chips - innovative tools for drug development and disease modeling, Lab Chip. 17 (2017) 3026–3036, https://doi.org/10.1039/c7lc00462a.

[100] A. Bal-Price, H.T. Hogberg, K.M. Crofton, M. Daneshian, R.E. FitzGerald, E. Fritsche, T. Heinonen, S.H. Bennekou, S. Klima, A.H. Piersma, M. Sachana, T.J. Shafer, A. Terron, F. Monnet-Tschudi, B. Viviani, T. Waldmann, R.H.S. Westerink, M.F. Wilks, H. Witters, M.-G. Zurich, M. Leist, Recommendation on test readiness criteria for new approach methods (NAM) in toxicology: exemplified for developmental neurotoxicity (DNT), ALTEX 35 (2018) 306–352. https://doi.org/10.14573/altex.1712081.

[101] A. Krebs, B.M.A. van Vugt-Lussenburg, T. Waldmann, W. Albrecht, J. Boei, B. ter Braak, M. Brajnik, T. Braunbeck, T. Brecklinghaus, F. Busquet, A. Dinnyes, J. Dokler, X. Dolde, T.E. Exner, C. Fisher, D. Fluri, A. Forsby, J.G. Hengstler, A.-K. Holzer, Z. Janstova, P. Jennings, J. Kisitu, J. Kobolak, M. Kumar, A. Limonciel, J. Lundqvist, B.

Mihalik, W. Moritz, G. Pallocca, A.P.C. Ulloa, M. Pastor, C. Rovida, U. Sarkans, J.P. Schimming, B.Z. Schmidt, R. Stöber, T. Strassfeld, B. van de Water, A. Wilmes, B. van der Burg, C.M. Verfaillie, R. von Hellfeld, H. Vrieling, N.G. Vrijenhoek, M. Leist, The EU-ToxRisk method documentation, data processing and chemical testing pipeline for the regulatory use of new approach methods, Arch Toxicol. 94 (2020) 2435–2461, https://doi.org/10.1007/s00204-020-02802-6.

[102] ICCVAM, ICCVAM Test Method Evaluation Report: In vitro cytotoxicity test methods for estimating starting doses for acute oral systemic toxicity testing. National Toxicology Program, Research Triangle Park, NC., (2006). https://ntp.niehs.nih.gov/iccvam/docs/acutetox_docs/brd_tmer/at-tmer-complete.pdf.

[103] NTP, Validation Study of In Vitro Cytotoxicity Test Methods, (2019). https://ntp.niehs.nih.gov/whatwestudy/niceatm/test-method-evaluations/acutesystemic-tox/in-vitro-validation/index.html (accessed January 21, 2022).

[104] ICCVAM, Background Review Document: In vitro cytotoxicity test methods for estimating acute oral systemic toxicity. Vol. 1., National Institute for Environmental Health Sciences, Research Triangle Park, NC, 2006. https://ntp.niehs.nih.gov/iccvam/docs/acutetox_docs/brd_tmer/brdvol1_nov2006.pdf.

[105] ICCVAM, Peer Review Panel Report: The use of in vitro basal cytotoxicity test methods for estimating starting doses for acute oral systemic toxicity testing., National Institute for Environmental Health Sciences, Research Triangle Park, NC, 2006. https://ntp.niehs.nih.gov/iccvam/docs/acutetox_docs/atpanelrpt06/atpanelrpt.pdf.

[106] OECD, Guidance document on using cytotoxicity tests to estimate starting doses for acute oral systemic toxicity tests, OECD Publishing, Paris, 2010. https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2010)20&doclanguage=en.

[107] A. Cherkasov, E.N. Muratov, D. Fourches, A. Varnek, I.I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y.C. Martin, R. Todeschini, V. Consonni, V.E. Kuz'min, R. Cramer, R. Benigni, C. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard, A. Tropsha, QSAR modeling: where have you been? Where are you going to? J. Med. Chem. 57 (12) (2014) 4977–5010.

[108] T.W. Schultz, A. Chapkanov, S. Kutsarova, O.G. Mekenyan, Assessment of uncertainty and credibility of predictions by the OECD QSAR Toolbox automated read-across workflow for predicting acute oral toxicity, Comput. Toxicol. 22 (2022) 100219, https://doi.org/10.1016/j.comtox.2022.100219.

[109] I. Shah, J. Liu, R.S. Judson, R.S. Thomas, G. Patlewicz, Systematically evaluating read-across prediction and performance using a local validity approach characterized by chemical structure and bioactivity information, Regul. Toxicol. Pharmacol. 79 (2016) 12–24, https://doi.org/10.1016/j.yrtph.2016.05.008.

[110] National Research Council, Toxicity Testing in the 21st Century: A Vision and a Strategy, The National Academies Press, Washington, DC, 2007. https://doi.org/ 10.17226/11970.

[111] A.P. Worth, G. Patlewicz, Integrated Approaches to Testing and Assessment, in: C. Eskes, M. Whelan (Eds.), Validation of Alternative Methods for Toxicity Testing, Springer International Publishing, Cham, 2016: pp. 317–342. https:// doi.org/10.1007/978-3-319-33826-2_13.

[112] S.W. Edwards, M. Nelms, V.K. Hench, J. Ponder, K. Sullivan, Mapping mechanistic pathways of acute oral systemic toxicity using chemical structure and bioactivity measurements, Front. Toxicol. 4 (2022) 824094, https://doi.org/10.3389/ftox.2022.824094.

[113] FAA, What are Dangerous Goods?, Federal Aviation Administration. (2022).https://www.faa.gov/hazmat/what_is_hazmat/ (accessed December 16, 2021).

[114] IATA, Dangerous Goods Regulations (DGR)., (2022). https://doi.org/www.iata.org/dgr.