Mind the Gaps: Prioritizing Activities to Meet Regulatory Needs for Acute Systemic Lethality

Abstract
Efforts are underway to develop and implement nonanimal approaches which can characterize acute systemic lethality. A workshop was held in October 2019 to discuss developments in the prediction of acute oral lethality for chemicals and mixtures, as well as progress and needs in the understanding and modeling of mechanisms of acute lethality. During the workshop, each speaker led the group through a series of charge questions to determine clear next steps to progress the aims of the workshop. Participants concluded that a variety of approaches will be needed and should be applied in a tiered fashion. Non-testing approaches, including waiving tests, computational models for single chemicals, and calculating the acute lethality of mixtures based on the LD_{50} values of mixture components, could be used for some assessments now, especially in the very toxic or non-toxic classification ranges. Agencies can develop policies indicating contexts under which mathematical approaches for mixtures assessment are acceptable; to expand applicability, poorly predicted mixtures should be examined to understand discrepancies and adapt the approach. Transparency and an understanding of the variability of *in vivo* approaches are crucial to facilitate regulatory application of new approaches. In a replacement strategy, mechanistically based *in vitro* or *in silico* models will be needed to support non-testing approaches especially for highly acutely toxic chemicals. The workshop discussed approaches that can be used in the immediate or near term for some applications and identified remaining actions needed to implement approaches to fully replace the use of animals for acute systemic toxicity testing.

1 Introduction

Information about a chemical's potential to cause acute systemic lethality in humans is commonly required by regulatory authorities. Acute systemic lethality tests may be conducted via the oral, dermal, or inhalation route, as well as intravenously for extracts of medical devices. The acute oral lethal dose 50 (LD_{50}) test – the focus of this workshop – was first introduced by the U.S. Food and Drug Administration in 1928 and has become entrenched in the regulatory system. Historically, the test was conducted to determine the dose of a substance that produces lethality in 50% of the animals tested; however, over the years, changes have been made to the test guidelines to reduce and refine animal use (Hamm et al., 2017).

Various international efforts have focused on replacing the use of animals in acute systemic lethality testing due to issues of reproducibility, human-relevance, and animal welfare. Progress has included the development of guidance on waiving these tests (EPA, 2020, 2016, 2012; OECD, 2017; PMRA, 2013), assessment and use of mathematical calculations to predict the toxicity of mixtures based on ingredients (UN, 2015), better understanding of human mechanisms-of-action, and the development of *in silico* models and *in vitro* tests.

This progress has come about due to strong leadership and cross-sector collaborations that have driven more human-relevant animal-free science. This has been particularly apparent within the Environmental Protection Agency (EPA), including an animal reduction goal set by the former EPA Office of Pesticide Program director[1], the amended Toxic Substances Control Act that minimizes testing on vertebrates[2], and the EPA Administrator's commitment to end tests on mammals by 2035[3]. Beyond EPA, commitments from other stakeholders to reduce and replace *in vivo* tests have also been made (ICCVAM, 2018). This leadership is essential for promoting the use, further development, validation, and regulatory acceptance of human-predictive approaches.

In 2015, a workshop co-organized by the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), the Physicians Committee for Responsible Medicine (PCRM), and the PETA International Science Consortium Ltd. brought together approximately 60 experts from US and international regulatory agencies, academia, nongovernmental organizations, and industry to summarize the state-of-the-science

---

Tab. 1: Workshop speakers, affiliations, and talk titles

Speaker	Title
Amy Clippinger, PETA Science Consortium International e.V.	Acute systemic toxicity testing: Past, present, and future
Warren Casey, NICEATM	Regulatory needs: Can existing data be used to derive acute lethality estimates without animal tests?
Raja Settivari, Corteva AgriScience	Calculating the toxicity of pesticide mixtures
Kamel Mansouri, NICEATM	Structure-based approaches to predicting acute toxicity: CATMoS
Kyle Glover, US Army	Integration of *in vitro* and *in silico* tools for acute toxicity prediction
Daniel Wilson, Dow Chemical	Where can we predict acute lethality with mechanistic *in vitro* models?
Sanjeeva J. Wijeyesakere, FMC	How should ADME be considered?
Steve Edwards, RTI International	What does the AOP framework contribute?

of nonanimal methods for acute systemic toxicity testing, identify gaps, and the steps needed to make progress (Hamm et al., 2017). Participants agreed that key to developing nonanimal approaches for regulatory purposes was to describe the acute systemic toxicity testing requirements of regulatory authorities, how the data are currently used, what information is actually needed, and what the path is to gain acceptance of new methods. A paper describing the needs and requirements in the U.S. has been published (Strickland et al., 2018) and another on global needs and requirements is under development. These documents provide an organized and transparent summary of the multifaceted regulatory framework to provide clarity on agency needs and expectations to help modelers focus their efforts.

There has been significant progress in the development of new approach methods for identifying acute systemic toxicants subsequent to the 2015 workshop. Accordingly, on October 30-31, 2019, the NICEATM and PCRM convened a workshop to reassess the state-of-the-science, remaining gaps, and priority actions. This workshop focused on acute oral lethality; targeted efforts to replace inhalation toxicity have been covered at other workshops (Clippinger et al., 2018), and several regulatory agencies have established policies to waive acute dermal toxicity testing (PMRA, 2017; EPA, 2020, 2016). Workshop participants discussed the following topics:

– acute oral toxicity data needs for US and international regulatory agencies for chemicals and mixtures
– variability of the *in vivo* rat acute oral toxicity test, the reference against which new approaches are compared
– usefulness and limitations of *in silico* models for acute oral toxicity
– determining the types of biological and mechanistic assays that would complement *in silico* model results
– challenges associated with testing mixtures
– usefulness and limitations of existing strategies for predicting acute oral toxicity of mixtures (e.g., the GHS Mixtures Equation)

Presentations and group discussions focused on issues specific to particular classes of chemicals and next steps (Tab. 1).

2 Understanding variability of the reference method

The extent to which the traditional *in vivo* test can reproducibly predict itself provides an important range of responses against which to determine whether a nonanimal method is sufficient for regulatory use. Ultimately, users need to know how large the confidence interval for any given model prediction can be (or, how far away from an expected prediction is acceptable). Accordingly, NICEATM is analyzing the variability of the *in vivo* acute oral test using LD_{50} values from approximately 5,000 chemicals that were tested more than once (Karmaus et al., in preparation). Paramount to this effort is extensive data curation that can partially be accomplished computationally but also requires expert judgement and thereby manual processes. This analysis will define a margin of uncertainty when using rat oral acute toxicity data to assess the performance of nonanimal methods and provide a reference dataset to ensure that appropriately representative LD_{50} data are routinely used for the development and validation of nonanimal models. Efforts are needed to understand the causes for such variability; for example, artifacts due to hydrolytic inactivation of chemically reactive compounds could potentially confound either *in vivo* or *in vitro* studies.

3 Evaluating mixtures

Acute toxicity assessments of mono-constituent substances could be evaluated directly with a variety of nonanimal approaches, and their single structure nature makes them amenable to computational modeling using cheminformatic approaches. Multi-constituent mixtures and formulations (which will be referred to simply as mixtures herein), each comprised of more than one substance, present a variety of challenges to the development of new approach methodologies for acute toxicity testing, including:

– Some mixture co-formulants could impact the solubility of individual components, thereby potentially altering the toxicity of the mixture.

– Individual components in proprietary mixtures may not be publicly available, and thus there can be no assessment of the components based on their individual toxicity. Without such information, modelers and test method developers outside of the entity that owns the mixture are not able to leverage that information.

– Even if all mixture components are known, their respective acute toxicities may be unknown, may be derived from read-across with structurally similar chemicals, or may vary where multiple studies are available.

In many cases, mixtures represent the bulk of the testing that is conducted for registration, thereby providing the greatest quantitative opportunity in terms of reducing and replacing the number of animals used in acute toxicity tests. For example, from 2013-2017, fewer than 5% of acute oral toxicity submissions to EPA were for registrations of new active ingredients (approximately 10-15 per year). In fact, most of the acute oral toxicity submissions (approximately 240 per year) were for new formulations based on registered active ingredients. Mixtures are also relevant to several other agencies, including the Consumer Product Safety Commission and the Department of Defense.

3.1 GHS mixtures equation to predict acute toxicity

The UN Globally Harmonized System (GHS) for Classification and Labeling provides a mathematical approach for estimating the toxicity of mixtures based on the combined toxicities of the individual components of the mixture. Although many countries have implemented the GHS classification criteria, there has yet to be worldwide adoption of the GHS Mixtures Equation despite the demonstration of its utility through retrospective evaluations (Corvaro et al., 2016; Van Cott et al., 2018). Using the combined datasets from these two publications, NICEATM conducted an evaluation of the GHS Mixtures Equation as it applies to EPA categories (Hamm et al., in preparation). While the combined dataset is skewed towards less toxic substances (e.g., over 90% of the mixtures included in (Corvaro et al., 2016) are EPA Category III or IV), the majority of discordance between the *in vivo* LD_{50} and the predicted LD_{50} using the GHS Mixtures Equation seems to be due to an under-prediction of EPA Category III as Category IV. However, most of the "underclassified" EPA Category III mixtures are based on *in vivo* LD_{50} values between 2000 to 5000 mg/kg. Since LD_{50} > 2000 mg/kg is not classified in most jurisdictions, these "underclassified" mixtures are likely of limited concern. Conversely, for most *in vivo* Category IV mixtures that are identified as "overclassified" based on the GHS Mixtures Equation, the calculated value is 2000 mg/kg < LD_{50} < 5000 mg/kg. Again, given the threshold for mixtures not requiring classification, such discordance does not present a concern.

In December 2016, EPA created a voluntary pilot program to evaluate the usefulness and acceptability of the GHS Mixtures Equation as it applies to EPA hazard classification and labeling for agrochemical formulations. In response, stakeholders submitted toxicity data paired with calculations done in accordance with the GHS Mixtures Equation to support evaluations of pesticide product formulations. A total of 491 combined datasets were submitted for acute oral lethality, most of which (444/491) were classified based on *in vivo* studies as EPA Category III or IV (the least toxic categories). Analyses are ongoing to further examine the utility of the GHS Mixtures Equation and to compare to the published results noted above (Hamm et al., in preparation). A definitive conclusion for the Equation's application to the most toxic categories (i.e., EPA Category I and II) will not likely be feasible given the small dataset of substances in these categories that were submitted to the pilot program and will require further data collection and analysis.

It should be noted that the GHS Mixtures Equation is a simple approach to calculate a weighted average LD_{50} based on concentration and measured LD_{50} of each mixture component. More complex mathematical approaches could be explored to improve performance if specifics were made available for the proprietary mixture components (e.g., identity, relative concentration, LD_{50}).

4 Computational approaches to predict acute oral toxicity

With an increasing number of diverse chemicals to assess for acute systemic toxicity potential, *in silico* models provide a potential approach to predict acute oral toxicity and bridge data gaps. While, in some cases, existing *in silico* models have been successful in predicting acute toxicity, additional work is needed to cover current regulatory needs. With the goal of developing such an approach, it is imperative to explore several questions, such as: Where do existing models fall short? How might biological information complement their utility? Can classes of chemicals/mechanisms be identified for which specific assay/model development is needed to predict acute toxicity?

Differences in the categories assigned based on *in vivo* study results and those predicted by *in silico* models could be due to a variety of reasons including (but not limited to): unequal distribution of toxicity potential of the model training set; errors in the dataset used to build the model; variability of the reference method; metabolism and detoxification processes; and reactive chemistries. To make *in silico* predictions for mixtures, considerations include the need to distinguish between formulations with similar compositions but varying forms (e.g., water-based and oil-based formulations) and the potential need to assess interactions between co-formulants. *In vitro* models could provide biological information (e.g., mechanisms, metabolism) to complement the utility of *in silico* models. Multiple *in vitro* assays are likely needed to cover the full spectrum of potential toxicity mechanisms, as it is not always known or obvious which mechanism drives high toxicity.

4.1 Consensus model to predict acute oral toxicity

NICEATM and the ICCVAM Acute Toxicity Workgroup organized an international collaborative project to develop *in silico* models for predicting acute oral toxicity of mono-constituent substances (Kleinstreuer et al., 2018). In total, 35 groups participated, submitting 139 predictive models built using a

dataset of 11,992 chemicals. Models were developed for five endpoints: LD_{50} value, EPA hazard categories, GHS hazard categories, very toxic ($LD_{50} < 50$ mg/kg), and non-toxic ($LD_{50} > 2000$ mg/kg). Predictions within the applicability domains of the submitted models were evaluated using external validation sets, then combined into consensus predictions for each endpoint, forming the Collaborative Acute Toxicity Modeling Suite CATMoS (Kleinstreuer et al., 2018). The resulting consensus predictions leverage the strengths and overcome the limitations of individual modeling approaches. The consensus predictions performed at least as well as the *in vivo* acute oral toxicity assay in terms of accuracy and reproducibility. The evaluation set balanced accuracy ranged from 74% to 84% for the four categorical endpoint predictions and an R^2 of 0.65 for LD_{50}. The CATMoS consensus model was implemented with an applicability domain assessment and accuracy estimates so that predictions can be generated for new chemical structures via the OPEn structure activity/property Relationship App (OPERA) predictive free and open-source tool and made publicly accessible via NTP's Integrated Chemical Environment (ICE) (Mansouri et al., 2018, submitted; Bell et al., 2020). As with any *in silico* tool, model outputs should be evaluated using expert judgement prior to implementation for regulatory use.

The Environmental Fate and Effects Division (EFED) of EPA's Office of Pesticide Programs uses LD_{50} data in a quantitative manner as part of their ecological risk assessments for new pesticide active ingredient registrations. EFED, in collaboration with the Humane Society of the United States and NICEATM, is assessing the model predictions for about 170 pesticides registered in the past approximately twenty years with respect to impacts on mammalian risk assessments were the model predictions to be used in place of the reported LD_{50} values. Model predictions are systematically substituted for LD_{50} values, followed by recalculation of mammalian risk quotients (RQs) (ratio of LD_{50} in mg/kg-bodyweight to predicted pesticide exposure in mg/kg-bodyweight under various product application scenarios and mammalian feeding scenarios). Recalculated RQs will be compared to EFED's Levels of Concern (LOC). Using the model prediction in place of the LD_{50} can have four results: 1) agreement that an LOC was triggered; 2) agreement that an LOC was not triggered; 3) an LOC was triggered with the model value but not with the LD_{50} value, thereby identifying a risk where previously there was none; 4) an LOC was triggered with the LD_{50} value, but not with the model value. Qualitative and quantitative assessments will be done to determine model performance and explore further any differences between the model prediction and *in vivo* LD_{50} values.

4.2 Profilers to predict acute toxicity

Computational approaches can be used to understand the relevance of mechanistic interactions for high acute toxicity and then build "mechanistic profilers" based on a combination of pre-defined SMARTS (SMILES [Simplified Molecular Input Line Entry System] Arbitrary Target Specification) filters as well as machine-learning approaches, including 2D structural scaffolding and fingerprinting or 3D protein-ligand docking. Profilers thus far

have been built based primarily on seven interactions: (1) facile chemical reactivity (Wilson et al., 2018; Wijeyesakere et al., 2018), (2) chelation, (3) non-specific hydrophobic interactions, (4) surfactancy, (5) denaturancy, (6) protonophoric activity, and (7) non-covalent interaction with some specific receptors, enzymes or organelles (Wilson et al., 2018; Wijeyesakere et al., 2018). The seventh type of interaction can include a vast number of interactions; thus, future efforts will be to collaboratively identify the remaining mechanistic targets for high toxicity and derive useful data for training the next generation of profilers (e.g., neuronal receptors and cardiac channels). Such profilers can provide both positive and negative predictivity when trained using HTS data and may have balanced accuracies exceeding 90% for non-promiscuous targets. The profilers can be applied in a toxicity endpoint-agnostic and route-agnostic manner. Their application to a curated database, such as the acute oral LD_{50} database housed in ICE, will allow for rapid visualization of which molecular initiating events may drive classification at low exposures, with the recognition that only a few distinct mechanisms drive high acute toxicity and classification. Database profiling also allows a better visualization of the proportion of potential mechanisms that still need to be identified for high acute toxicity for prioritization of further research (Wijeyesakere et al., 2018; Wilson et al., 2018).

The profilers can be used as covariates in QSAR models for systemic toxicity and should be impactful as long as the given mechanism drives toxicity for that endpoint. Such profilers may suggest a testable hypothesis (e.g., whether a chemically reactive moiety drives the acute lethality or justifies whether to use a compound for read-across). For toxicity databases that might lack sufficient data to derive robust machine-learning models, they could also be leveraged to align a given compound to a specific *in vitro* test method or to provide insight into the performance of *in vitro* methods in general. Application of these methods to allow for identification of potential molecular initiating events should have broad application in the realm of predictive toxicology and for reductions in animal use. Their contributions in a weight-of-evidence approach will better enable trust in nonanimal methodologies, including the use of read-across, use of *in vitro* data, and identifying spurious outliers.

5 Consideration of mechanisms of action

Current regulatory needs and uses for acute lethality data do not typically include or reflect information on the mechanisms of action of the chemicals under scrutiny (Strickland et al., 2018). The performance of nonanimal methods for predicting acute systemic toxicity may be improved by considering specific mechanisms that drive high toxicity at low exposures (Wijeyesakere et al., 2018; Wilson et al., 2018). Understanding mechanisms of acutely toxic chemicals could also help to identify areas where assay development is needed, prioritize specific nonanimal methods for optimization, improve read-across, support the relevance of specific nonanimal methods to the endpoint, and impact emergency response measures. Further, this information can help to

Tab. 2: Mechanisms of acute lethality for which work is needed to further develop assays or models and prototypic substances known to act via those mechanisms

Mechanism	Example stressors
Voltage-gated channel (Na^+, K^+, Ca^{2+}) interaction	Marine toxins (Al-Sabi et al., 2006), pyrethroids, DDT
Dopaminergic	Ergot alkaloids
Histaminergic	Diphenhydramine, doxylamine
Protein synthesis inhibition	Ricin (Sowa-Rogozinska et al., 2019)
Dihydrofolate reductase inhibition	Methotrexate, pyrimethamine
TRPA1 interaction	Allylisothiocyanate, other isocyanates
Adrenergic interaction	Propranolol, phentolamine
Opioid receptor interaction	Fentanyl, morphine
Tubulin binding	Colchicine
Norepinephrine reuptake inhibition	Desvenlafaxine, levomilnacipran
NMDA receptor inhibition	Ammonium
Cardiac channel blocking	Diltiazem
Heme biosynthesis inhibitors	Lead, arsenic
Serotonin reuptake inhibitors	Citalopram (Isbister et al., 2004)

address some limitations of existing computational models, such as the propensity of datasets for model building that are skewed towards non-toxic compounds, thereby precluding the development of models for the most lethal compounds.

One major challenge in identifying relevant mechanisms of acute toxicity is that only limited information is collected from the acute *in vivo* tests, such as clinical signs, time-to-death, and gross pathology findings – and even these data are rarely made available in curated publicly available datasets. Mechanistic information collected from repeat-dose tests could be considered applicable to acute exposures although the number of mechanisms responsible for high acute toxicity by the oral route is limited. These often include obvious manifestations of well-characterized biological mechanisms of intended end-use application, such as anticoagulation by rodenticides, cholinesterase inhibition by organophosphate pesticides, and voltage-gated channel blocking by pyrethroids and pesticides such as dichlorodiphenyltrichloroethane (DDT). As well, some pharmacologic agents for pain relief have well-characterized biology and much human data from inadvertent overdosing, such as the current world-wide opioid crisis where the target is the pharmacologic mu-opioid receptor. Efforts are underway to comprehensively identify mechanisms important for acute lethality (Hamm et al., 2017; Wijeyesakere et al., 2018; Wilson et al., 2018; Prieto et al., 2019). Workshop discussions focused on determining where efforts need to be expended in order to efficiently increase coverage of acute lethality mechanisms.

As mentioned above, mechanistic interactions relevant for toxicity are one of at least seven types (covalent chemical reaction chemistry, chelants, non-specific hydrophobic interactions, sur-

factants, denaturants, protonophores, and non-covalent specific interaction with receptors, enzymes or organelles). Some initiating interactions, such as chelation, are thought to rarely manifest in high acute toxicity by the oral route; however, they can lead to secondary sequelae, such as precipitation of less soluble chelant-ligand complexes in the kidney or rendering essential elements such as zinc less bioavailable. Existing *in vitro* models (such as basal cytotoxicity or certain receptor binding assays) and *in silico* models may cover a large portion of the relevant chemical space. Recent implementation of mechanistic profilers allowed for retrospective identification of drug products that had been withdrawn from the market due to idiosyncratic acute liver injury (Wijeyesakere et al., 2019). Implementation of computational mechanistic profiling on "cytotoxicity" databases allows (1) quality inspection of the robustness of given assays, (2) subcategorization of types of cytotoxicity, and (3) regional mechanistic extrapolation to *in vivo* data for building regional models.

Work is needed to identify and develop profilers for the remaining mechanisms that drive high acute oral toxicity. Additionally, since preliminary screening has pointed to potential differences between mechanistic drivers of classification for acute oral versus inhalation toxicity, development of inhalation-enriched profilers may help establish integrated predictive approaches for inhalation toxicity, where the availability of fewer *in vivo* training data may necessitate combined application of *in silico* and *in vitro* models.

Work is also needed to develop assays to detect chemicals that are metabolized from parent compounds into more toxic species, including those that transform into cyanide, hydrogen sulfate, carbon monoxide, or nitrite. Metabolic transformation of com-

pounds that may inhibit aconitase can drive toxicity by inactivation of the Krebs cycle. Cytochrome P450-mediated metabolism of phosphothionates and some other sulfur-containing organophosphates can drive acute toxicity *in vivo*, but this is currently challenging to represent *in vitro*. Some compounds chelate only after enzymatic hydrolysis of ester or amide bonds, thus also requiring metabolism for acute toxicity.

Mechanisms of acute oral lethality for which work is needed to further develop assays or models include voltage-gated channel blockers, adrenergics, protein synthesis inhibitors, opioid receptor binders, tubulin binders, norepinephrine reuptake inhibitors, NMDA receptor inhibitors, heme biosynthesis inhibitors, and serotonin reuptake inhibitors (Tab. 2). A model for TRPV1 (transient receptor potential cation channel, subfamily V, member 1) was recently finished and is proving informative, especially for inhalation applications.

While investigation of acute toxicity mechanisms can be done independent of exposure route using certain assumptions, there can be notable differences in the responsible mechanisms between oral and inhaled exposures. This leads to the need for a combination of approaches, as determination of the most likely route of exposure may drive the selection of appropriate *in vitro* toxicity models. This information also helps to focus and prioritize efforts to develop new models adapted to the prediction of toxicity based on the relevant route of human exposure. The current emphasis is on the mechanisms that drive high toxicity and classification because of the responsibility to not miss such compounds and because there are far fewer mechanisms.

6 Case studies demonstrate need for a toolbox approach

Two case studies demonstrating the application of *in vitro* and *in silico* tools to characterize the acute lethality potential of chemicals relevant to the Department of Defense were presented for discussion by the workshop participants. The U.S. Army Combat Capabilities Development Command Chemical Biological Center (CBC) has the responsibility of rapidly identifying potential chemical threats and determining how to address them. As traditional *in vivo* methods cannot meet this demand, the CBC is creating a 4-tiered predictive toxicology toolbox to create rapid predictions of preliminary human estimates of toxicity and mechanism-informed toxicity extrapolations (Fig. 1).

The first case study aimed to develop a toolbox for 54 reactive chemical weapon precursor compounds, comprising a range of toxicities and chemical structure groups. Following the tiered approach, publicly available data (e.g., EPA ToxCast data) were gathered, and the chemicals were grouped by structure, leading to an assessment of the potential for read-across. Machine learning tools, including variable nearest neighbor (vNN), random forest, and artificial neural network (aNN) approaches, were used to fill data gaps for acute lethality values and then both measured

and predicted values were combined with physiologically based pharmacokinetic (PBPK) modeling, route-to-route extrapolation, and bioavailability based on Cmax oral and IV values to derive a 2-minute human 50% lethal concentration (LCt_{50}). These tools provide a preliminary estimate of potential human lethal toxicity and can be extremely useful to derive these predictions quickly but come with uncertainties that are important to understand and address. For example, the datasets used to create the QSAR models contain more structural diversity and representation at the less toxic end of the spectrum, which increases the likelihood of misalignment of predictions with LD_{50} values. The CBC found vNN to be the best method for addressing this limitation. Continued work will include refining chemical-specific PBPK models.

The second case study aimed to develop a model for predicting the potency of a list of opioid compounds. Nonanimal approaches were particularly appropriate for this case study as some animals are less sensitive to opioids than humans. The models explored included a drug discovery database called ToxTool, a receptor docking model, an *in vitro* receptor binding and functional assays as well as a limited rabbit study to generate PBPK values to refine *in vitro* potency predictions. Although human data are minimal, one case report aligned well with the potency prediction for the opioid carfentanil.

Both case studies demonstrated the importance of understanding the mechanism of action of the chemicals, the need to consider absorption, distribution, metabolism and excretion (ADME) of the chemical in humans in order to make predictions about the acute lethality potential to humans, and the utility of combining *in silico* and *in vitro* approaches in an integrated approach.

Other cases where metabolic processes affect the acute toxicity of compounds include cytochrome P450 related metabolism – important for understanding the inhibition potency of organophosphate compounds – and ester hydrolysis. Ester content correlates with GHS toxicity category when looking at rat acute oral toxicity data (i.e., the higher toxicity GHS 1-3 classes are less enriched in esters suggesting that enzymatic hydrolysis of parent compounds may often be a detoxification event (Wilson and Wijeyesakere, in preparation)). Other key metabolic processes include reductases, phase II conjugation, oxidative pathways, and H_2S releasers, among others. There is a need to catalog these processes and link them to known acutely toxic chemicals.

Workshop participants agreed that, while there are some assays available to assess metabolic processes, including *in vitro* hepatocyte and gut lining models and QSAR predictors, improved capacity for assessing potential metabolic processes is needed. Some publicly available tools, including high-throughput toxicokinetics and the OPERA pKa model (Mansouri et al., 2019), are potential areas for further case study development.

Another tool that can be more widely used for predicting acute lethality is the adverse outcome pathway (AOP) framework (Ankley et al., 2010). There are a few AOPs in the OECD AOP Wiki[4] that are directly relevant to acute lethality (Prieto et al., 2019), and the modular nature of the AOP Wiki elements and the ability to

---
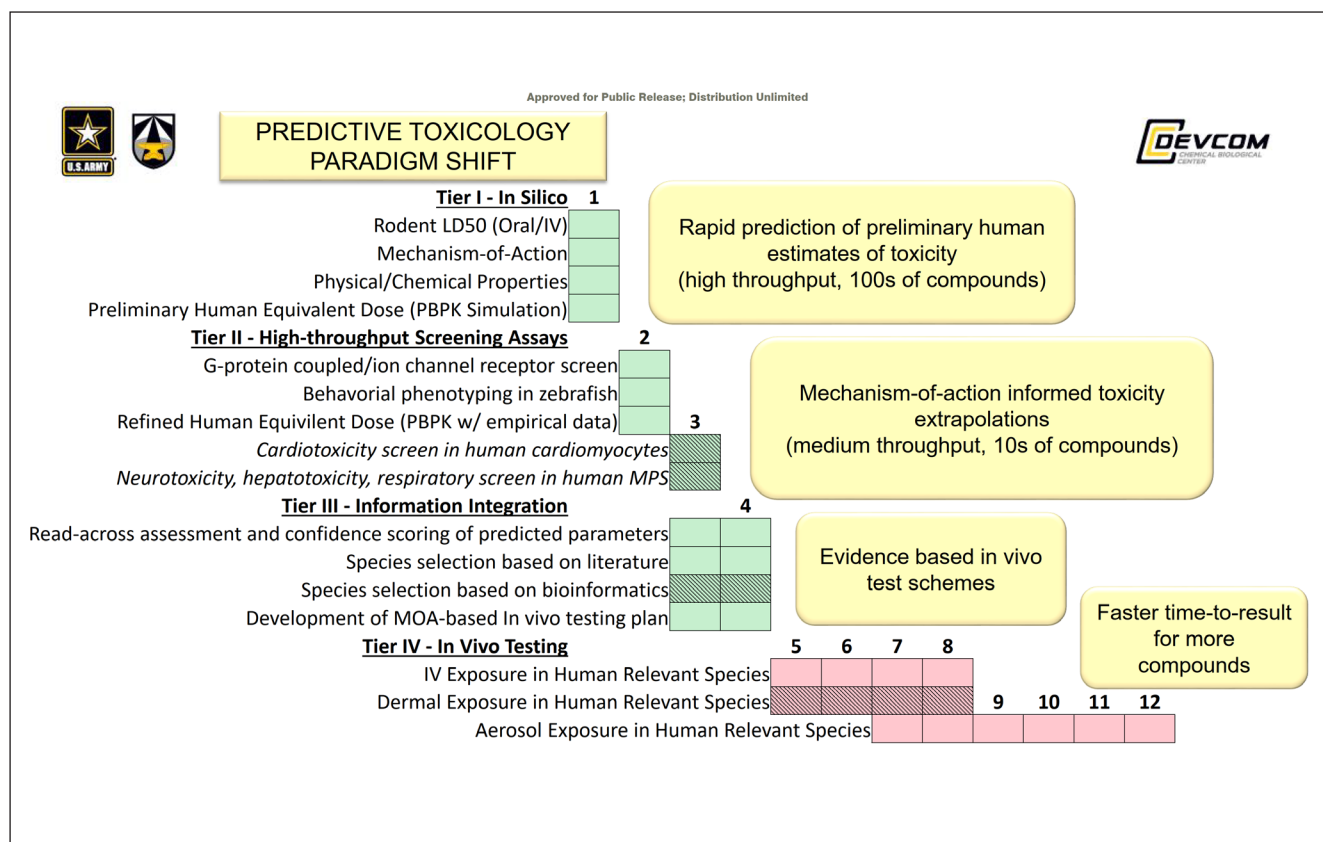
[4] Online at www.aopwiki.org

Fig. 1: Four-tiered predictive toxicology toolbox
A four-tiered predictive toxicology toolbox uses a combination of physical-chemical properties, computational models, and *in vitro* assays, with *in vivo* testing considered as a final tier if sufficient information cannot be otherwise obtained. This approach identifies relevant toxicities and provides more information faster than a traditional *in vivo* approach. Patterned blocks are optional depending on the information needed. PBPK, physiologically based pharmacokinetic; MPS, microphysiological system; MOA, mode of action

source information from a diverse set of experts worldwide make the framework ideally suited to organize information about mechanisms of acute lethality, given the diverse set of relevant potential toxicities and mechanisms.

The workshop discussion identified several key benefits of the AOP framework in this context. First, AOPs can support the regulatory acceptance of computational or *in vitro* tools by providing toxicological relevance and supporting their use in place of *in vivo* assays. This information can also reduce testing by helping to set aside or prioritize chemicals that do or do not act via certain mechanisms. Finally, having a more comprehensive view of interrelated pathways of acute toxicity could help to identify shared upstream key events that should be prioritized for model development or use, including models to predict toxicity of mixtures.

It was agreed that resources should not be spent on developing lengthy, detailed AOPs for each potential mechanism of acute lethality. Rather, efforts should build on what has already been done to compile relevant mechanisms. This mechanistic information can be added to the AOP Wiki, creating basic AOPs that can be filled in more completely as needed, depending on the state of the science or the particular application.

7 Follow-up activities

There is a need for repeatable and reproducible approaches that predict acute oral lethality in humans as well as or better than the currently used animal tests. Developing nonanimal approaches to replace acute lethality testing involves a variety of practical and scientific considerations. It is fundamentally important to distinguish between the information that is obtained from the animal test versus the information this is actually needed for regulators to protect human health and, consequently, what is needed from a new approach. It is also essential to understand the human relevance of information obtained from the animal test as well as the variability associated with it. An analysis of the variability of the *in vivo* data can provide context to the predictive capacity of the model or approach being considered and help regulators interpret the resulting information based on how the information is used.

The diversity of acutely toxic chemicals and mechanisms via which they act, combined with a large number of regulatory needs and uses for these data, compels consideration of a variety of methods and approaches as well as policy changes to facilitate their implementation. A tiered and flexible approach that

takes advantage of currently available tools while adapting as new tools are developed is needed.

Currently, a number of approaches can be used to reduce animal testing for acute oral lethality. For example, consensus models built on large curated datasets (e.g., CATMoS) may cover a large chemical space. Limitations of these models may be addressed by using specific mechanistic models, *in vitro* assays, expert-driven read-across, and ADME models and assays. There is a critical need to train subject-matter experts in the implementation of alternative models for acute mammalian toxicity.

Regulatory acceptance of nonanimal methods should be pursued in close collaboration with the relevant regulatory agencies, as the transparency, predictivity, and ease of use of models or data may affect whether an approach is accepted. For example, the use of *in vitro* assays or *in silico* models with proprietary elements or data sets may be accepted as long as a regulatory agency can evaluate the model elements and performance. Representation of these characteristics in a document, such as a QSAR Model Reporting Form (QMRF), can help to facilitate this evaluation.

One continued need is the determination of acute lethality for mixtures, such as pesticide formulations. Mathematical approaches, such as the GHS Mixtures Equation, can provide a weight-of-evidence classification for these mixtures that could be improved by incorporating information about formulation components. Work in progress could ultimately lead to the conclusion that such an approach reduces *in vivo* testing by reliably identifying negative outcomes, but a definitive conclusion requires that additional toxic compounds be compared (Hamm et al., in preparation).

To facilitate the use and regulatory acceptance of mathematical approaches for mixtures, the following follow-up activities were identified during the workshop:

– Publish agency policies indicating contexts under which mathematical approaches are acceptable.
– Examine why mathematical approaches do not correlate with *in vivo* LD_{50} values for some mixtures. A detailed analysis requires companies to confidentially share formulation information with agencies, such as the amount(s) and type(s) of solvent(s) and any surfactants contained in formulations.
– Collect additional information needed to conduct further analysis of the mathematical approaches. For example, while the EPA OPP pilot program collected substantial data on EPA category III and IV formulations, it is interested in expanding its dataset for EPA category I and II pesticides as well as for antimicrobial products.
– Evaluate evidence that might suggest potential mixture interactions of concern.
– Develop *in vitro* assays to use when data on components are not available or mathematical approaches are not applicable. This may be particularly useful for non-pesticide mixtures, such as consumer products or device extracts.

To make progress in developing and applying more mechanistic information and models, the following actions were identified during the workshop:

– Catalog and organize key metabolic processes and match these with available metabolically active *in vitro* assays to see where the gaps in development remain.
– Codify the mechanistic information collated in the literature and during the workshop into the AOP Wiki to share and analyze what is available and encourage further development by the scientific community.
– Encourage public hosting (within an open access resource such as ICE) of ADME/PBPK tools and models covering specific mechanistic domains to allow use by others and encourage further development.
– Prioritize development of new mechanistic models and assays based on regulatory needs.
– Conduct and publish proof-of-concept case studies to show the utility of and gain confidence in *in silico* and *in vitro* methods for predicting acute oral lethality.
– Organize hands-on training with computational tools to increase familiarity on how to use and interpret data from these models.
– Discuss needs with relevant stakeholders, including industry and regulatory partners, to facilitate implementation and acceptance of nonanimal approaches.

Gaining regulatory acceptance for nonanimal approaches to predict acute oral lethality requires a multi-pronged approach, including strong leadership from government agencies and stakeholder commitment. The discussions during this workshop provide important insights into remaining gaps and a roadmap to acceptance.

References

Al-Sabi, A., McArthur, J., Ostroumov, V. et al. (2006). Marine toxins that target voltage-gated sodium channels. *Marine Drugs 4*, 157-192. doi:10.3390/md403157

Ankley, G. T., Bennett, R. S., Erickson, R. J. et al. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem 29*, 730-741. doi:10.1002/etc.34

Bell, S., Abedini, J., Ceger, P. et al. (2020). An integrated chemical environment with tools for chemical safety testing. *Toxicol In Vitro 67*, 104916. doi:10.1016/j.tiv.2020.104916

Clippinger, A. J., Allen, D., Behrsing, H. et al. (2018). Pathway-based predictive approaches for non-animal assessment of acute inhalation toxicity. *Toxicol In Vitro 52*, 131-145. doi:10.1016/j.tiv.2018.06.009

Corvaro, M., Gehen, S., Andrews, K. et al. (2016). GHS additivity formula: A true replacement method for acute systemic toxicity testing of agrochemical formulations. *Regul Toxicol Pharmacol 82*, 99-110. doi:10.1016/j.yrtph.2016.10.007

EPA – Environmental Protection Agency (2012). Guidance for Waiving or Bridging of Mammalian Acute Toxicity Tests for Pesticides and Pesticide Products (Acute Oral, Acute Dermal, Acute Inhalation, Primary Eye, PrimaryDermal, and Dermal Sensitization). https://www.epa.gov/sites/production/files/documents/acute-data-waiver-guidance.pdf

EPA (2016). Guidance for Waiving Acute Dermal Toxicity Tests for Pesticide Formulations & Supporting Retrospective Analysis. https://bit.ly/3bSAY1W

EPA (2020). Public Review Draft - Guidance for Waiving Acute Dermal Toxicity Tests for Pesticide Technical Chemicals & Supporting Retrospective Analysis, September 2020. https://bit.ly/2P4LYQI

Hamm, J., Sullivan, K., Clippinger, A. J. et al. (2017). Alternative approaches for identifying acute systemic toxicity: Moving from research to regulatory testing. *Toxicol In Vitro 41*, 245-259. doi:10.1016/j.tiv.2017.01.004

ICCVAM (2018). A strategic roadmap for establishing new approaches to evaluate the safety of chemicals and medical products in the United States. doi:10.22427/NTP-ICCVAM-ROADMAP2018

Isbister, G. K., Bowe, S. J., Dawson, A. et al. (2004). Relative toxicity of selective serotonin reuptake inhibitors (SSRIs) in overdose. *J Toxicol Clin Toxicol 42*, 277-285. doi:10.1081/clt-120037428

Kleinstreuer, N. C., Karmaus, A., Mansouri, K. et al. (2018). Predictive models for acute oral systemic toxicity: A workshop to bridge the gap from research to regulation. *Comput Toxicol 8*, 21-24. doi:10.1016/j.comtox.2018.08.002

Mansouri, K., Grulke, C. M., Judson, R. S. et al. (2018). Opera models for predicting physicochemical properties and environmental fate endpoints. *J Cheminform 10*, 10. doi:10.1186/s13321-018-0263-1

Mansouri, K., Cariello, N. F., Korotcov, A. et al. (2019). Open-source QSAR models for pKa prediction using multiple machine learning approaches. *J Cheminform 11*, 60. doi:10.1186/s13321-019-0384-1

Mansouri, K., Karmaus A., Fitzpatrick J. et al. (submitted). CATMoS: Collaborative acute toxicity modeling suite. *Environ Health Perspective*s.

OECD – Organisation for Economic Cooperation and Development (2017). Guidance Document on Considerations for Waiving or Bridging of Mammalian Acute Toxicity Tests. *Series on Testing and Assessment, No. 237*. OECD Publishing, Paris. doi:10.1787/9789264274754-en

PMRA Health Canada (2013). Guidance for Waiving or Bridging of Mammalian Acute Toxicity Tests for Pesticides. https://bit.ly/3cBmvXo

PMRA Health Canada (2017). Science Policy Note SPN2017-03, Acute Dermal Toxicity Study Waiver. https://bit.ly/2OZa9QT

Prieto, P., Graepel, R., Gerloff, K. et al. (2019). Investigating cell type specific mechanisms contributing to acute oral toxicity. *ALTEX 36*, 39-64. doi:10.14573/altex.1805181

Sowa-Rogozinska, N., Sominka, H., Nowakowska-Golacka, J. et al. (2019). Intracellular transport and cytotoxicity of the protein toxin ricin. *Toxins (Basel) 11*, 350. doi:10.3390/toxins11060350

Strickland, J., Clippinger, A. J., Brown, J. et al. (2018). Status of acute systemic toxicity testing requirements and data uses by U.S. regulatory agencies. *Regul Toxicol Pharmacol 94*, 183-196. doi:10.1016/j.yrtph.2018.01.022

UN – United Nations (2015). Globally harmonized system of classification and labelling of chemicals (GHS) – Sixth revised edition. doi:10.18356/1e984724-en

Van Cott, A., Hastings, C. E., Landsiedel, R. et al. (2018). GHS additivity formula: Can it predict the acute systemic toxicity of agrochemical formulations that contain acutely toxic ingredients? *Regul Toxicol Pharmacol 92*, 407-419. doi:10.1016/j.yrtph.2017.12.024

Wijeyesakere, S. A., Wilson, D., Settivari, R. S. et al. (2018). Development of a profiler for facile chemical reactivity using the open-source Konstanz information miner. *Appl In Vitro Toxicol 4*, 202-213. doi:10.1089/aivt.2017.0040

Wijeyesakere, S. A., Wilson, D., Auernhammer, T. R. et al. (2019). Hybrid machine-learning/smarts profiling model for mitochondrial inhibition. *Appl In Vitro Toxicol 5*, 196-204. doi:10.1089/aivt.2019.0010

Wilson, D., Wijeyesakere, S. A., Parks, A. K. et al. (2018). Profiling acute oral and inhalation toxicity data using a computational workflow to screen for facile chemical reactivity. *Appl In Vitro Toxicol 4*, 214-219. doi:10.1089/aivt.2017.0041

Acknowledgements

Kristie M. Sullivan[1], David G. Allen[2], Amy J. Clippinger[3], Daniel M. Wilson[4], Stephen W. Edwards[5], Kyle Glover[6], Kamel Mansouri[2], Raja Settivari[7], Sanjeeva J. Wijeyesakere[8] and Warren Casey[9]

[1]Physicians Committee for Responsible Medicine, Washington, DC, USA; [2]Integrated Laboratory Systems, Research Triangle Park, NC, USA; [3]PETA Science Consortium International e.V., Stuttgart, Germany; [4]The Dow Chemical Company, Midland, MI, USA; [5]RTI International, Research Triangle Park, NC, USA; [6]US Army Combat Capabilities Development Command, Chemical Biological Center, APG, MD, USA; [7]Corteva Agriscience, Newark, DE, USA; [8]FMC Corporation, Newark, DE, USA; [9]National Toxicology Program, Research Triangle Park, NC, USA

(ksullivan@pcrm.org)