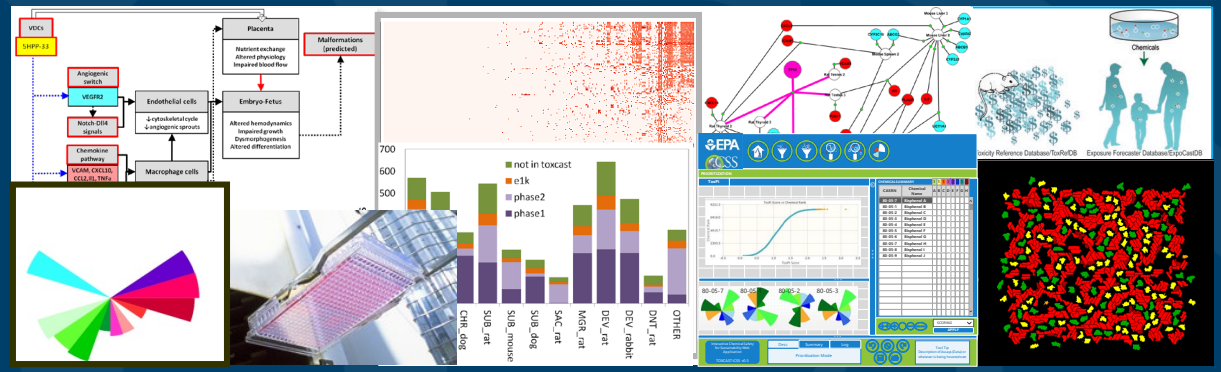
Implementation of Nonanimal Approaches for Acute Systemic Toxicity



Grace Patlewicz
National Center for Computational Toxicology (NCCT), US EPA

Presenting as co-chair & member of the ICCVAM Acute Toxicity Work Group (ATWG)

Acknowledgements

- **NICEATM**
- Nicole Kleinstreuer**
- **ILS**
- Agnes Karmaus**
- Kamel Mansouri**
- Dave Allen
- **EPA-NCCT**
- Jeremy Fitzpatrick**
- Prachi Pradeep**
- George Helman
- Imran Shah

**Also provided all slide materials for this presentation

https://ntp.niehs.nih.gov/go/natl-strategy

Outline

- ICCVAM Workgroup on Acute Toxicity – Charges & Scope
- 'Highs and Lows' in developing new and evaluating existing non-animal alternative approaches to acute toxicity testing
- Summary remarks

ICCVAM

- Interagency Coordinating Committee for the Validation of Alternative Methods
- H.R. 4281 (106th): ICCVAM Authorization Act of 2000
- To establish, wherever feasible, guidelines, recommendations, and regulations that promote the regulatory acceptance of new and revised toxicological tests that protect human and animal health and the environment while reducing, refining, or replacing animal tests and ensuring human safety and product effectiveness.

7 Regulatory Agencies

Consumer Product Safety Commission
Department of Agriculture
Department of the Interior
Department of Transportation
Environmental Protection Agency
Food and Drug Administration
Occupational Safety and Health Administration

9 Research Agencies

Agency for Toxic Substances and Disease Registry
National Institute for Occupational Safety and Health
National Cancer Institute
National Institute of Environmental Health Sciences
National Library of Medicine
National Institutes of Health
Department of Defense
Department of Energy
National Institute of Standards and Technology

- Other participants include: NCATS , Tox21 Representatives

ATWG Acute Toxicity Implementation Plan

- Coordinate activities via ICCVAM Workgroups
- Draft a scoping document to identify U.S. agency requirements, needs, and decision contexts for acute toxicity data
- Coordinate efforts with stakeholders
- Identify, acquire, and curate high quality data from reference test methods
- Identify and evaluate non-animal alternative approaches to acute toxicity testing
- Gain regulatory acceptance and facilitate use of non-animal approaches

ATWG Acute Toxicity Implementation Plan

- Coordinate activities via ICCVAM Workgroups
- Draft a scoping document to identify U.S. agency requirements, needs, and decision contexts for acute toxicity data
- Coordinate efforts with stakeholders
- Identify, acquire, and curate high quality data from reference test methods
- Identify and evaluate non-animal alternative approaches to acute toxicity testing
- Gain regulatory acceptance and facilitate use of non-animal approaches

ATWG Acute Toxicity Implementation Plan

- Coordinate activities via ICCVAM Workgroups
- Draft a scoping document to identify U.S. agency requirements, needs, and decision contexts for acute toxicity data
- Coordinate efforts with stakeholders
- Identify, acquire, and curate high quality data from reference test methods
- Identify and evaluate non-animal alternative approaches to acute toxicity testing
- Gain regulatory acceptance and facilitate use of non-animal approaches

Identify U.S. agency requirements, needs, and decision contexts for acute toxicity data

Hazard

- I (≤ 50mg/kg)
- II (>50 ≤ 500mg/kg)
- III (>500 ≤ 5000mg/kg)
- IV (>5000mg/kg)

Hazard

- Highly toxic (≤50mg/kg)
- Toxic (>50-5000mg/kg)

Packing Group

- I (≤ 5mg/kg)
- II (>5 ≤ 50mg/kg)
- III (>50 ≤ 300mg/kg)
- IV (>300 ≤ 2000mg/kg)

Hazard

GHS

Regulatory Toxicology and Pharmacology 94 (2018) 183–196

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

Status of acute systemic toxicity testing requirements and data uses by U.S. regulatory agencies

Judy Strickland[a,*], Amy J. Clippinger[b], Jeffrey Brown[b], David Allen[a], Abigail Jacobs[c,1], Joanna Matheson[d], Anna Lowit[e], Emily N. Reinke[f], Mark S. Johnson[f], Michael J. Quinn Jr.[f], David Mattie[g], Suzanne C. Fitzpatrick[h], Surender Ahir[i], Nicole Kleinstreuer[j], Warren Casey[j]

ARTICLE INFO

ABSTRACT

Acute systemic toxicity data are used by a number of U.S. federal agencies, most commonly for hazard classification and labeling and/or risk assessment for acute chemical exposures. To identify opportunities for the implementation of non-animal approaches to produce these data, the regulatory needs and uses for acute systemic toxicity information must first be clarified. Thus, we reviewed acute systemic toxicity testing requirements for six U.S. agencies (Consumer Product Safety Commission, Department of Defense, Department of Transportation, Environmental Protection Agency, Food and Drug Administration, Occupational Safety and Health Administration) and noted whether there is flexibility in satisfying data needs with methods that replace or reduce animal use. Understanding the current regulatory use and acceptance of non-animal data is a necessary starting point for future method development, optimization, and validation efforts. The current review will inform the development of a national strategy and roadmap for implementing non-animal approaches to assess potential hazards associated with acute exposures to industrial chemicals and medical products. The Acute Toxicity Workgroup of the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM), U.S. agencies, non-governmental organizations, and other stakeholders will work to execute this strategy.

ATWG Acute Toxicity Implementation Plan

- Coordinate activities via ICCVAM Workgroups
- Draft a scoping document to identify U.S. agency requirements, needs, and decision contexts for acute toxicity data
- Coordinate efforts with stakeholders
- Identify, acquire, and curate high quality data from reference test methods ..and evaluate the variability of the data..
- Identify and evaluate non-animal alternative approaches to acute toxicity testing
- Gain regulatory acceptance and facilitate use of non-animal approaches

Rat oral acute toxicity LD50 Database

- Mined and merged multiple existing resources containing rat oral acute toxicity LD50 data (collaboration between NICEATM & NCCT)
- Identify transcription errors (e.g. 20005000 mg/kg)
- Manual curation of highly variable chemicals; identify source data
- Often (typically) meta data not available for vast majority of the substances collected
- Explore the variability of the data - representative LD50, variability across hazard categories

Data source	Number of LD50 values	Number of unique chemicals
ECHA ChemProp	5,533	2,136
NLM HSDB	3,981	2,205
JRC AcutoxBase	637	138
NLM ChemIDplus	13,072	12,977
NICEATM PAI	364	293
OECD eChemPortal	10,119	2,290

Total:
34,511 LD50 values
16,307 chemicals

↓ Identify unique data in mg/kg

21,210 LD50 values
15,698 chemicals

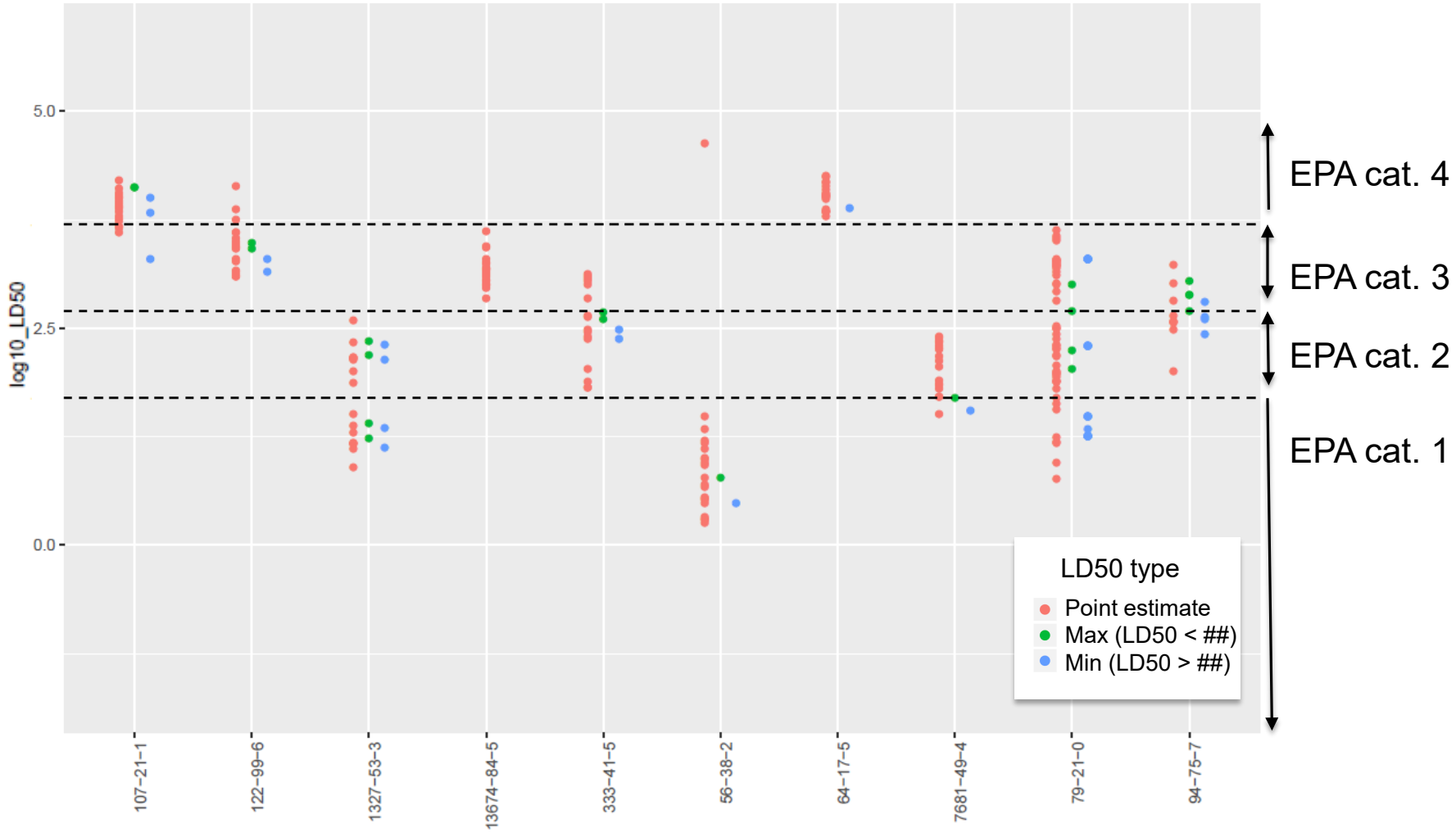
Acute Oral LD50 Dataset Replicate Inventory

- 13,339 chemicals with one LD50 value
- 2,349 chemicals with ≥2 LD50 values
- 1,120 chemicals with ≥3 LD50 values
- 609 chemicals with ≥4 LD50 values
- 347 chemicals with ≥5 LD50 values

Orders of magnitude for LD50s	Number of chemicals
0	546 (49%)
1	519 (46%)
2	39 (3%)
3	8 (0.7%)
4	8 (0.7%)

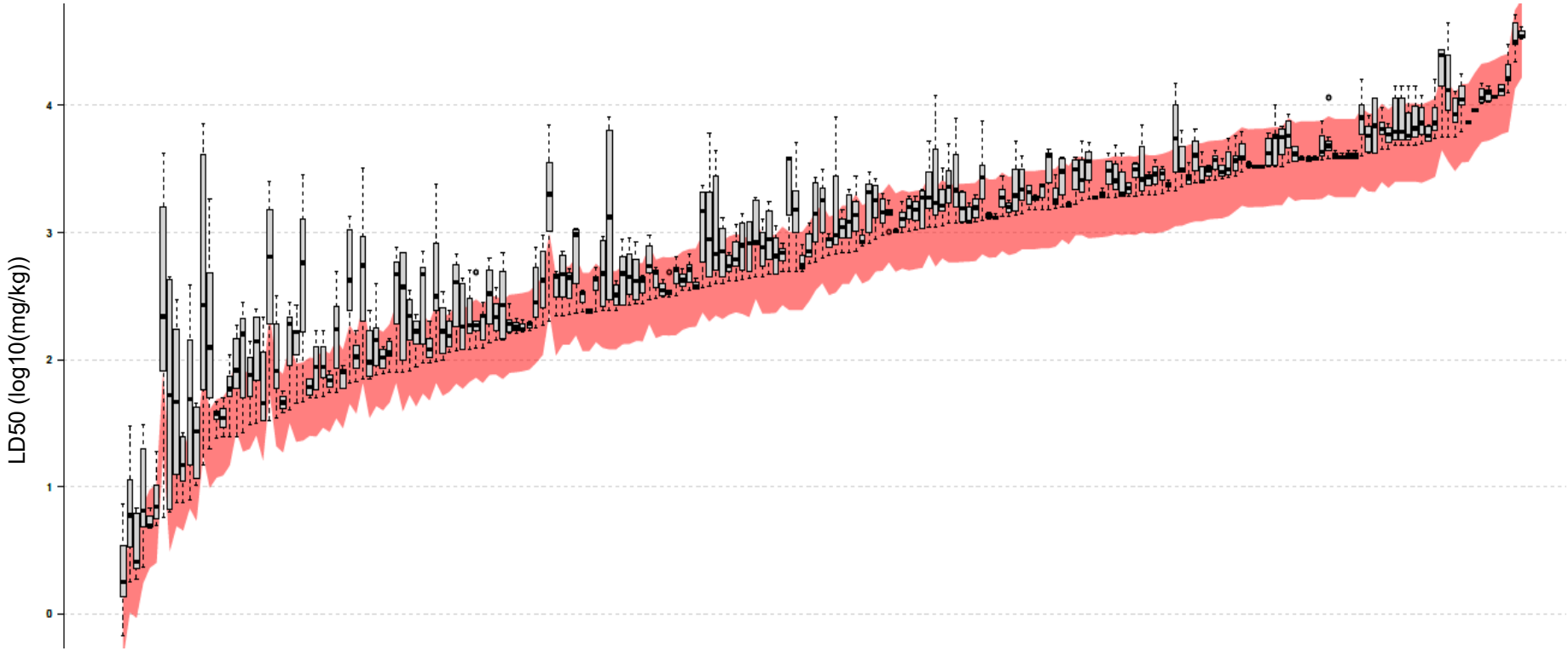
Impact on Hazard Categorization

Example: EPA Classification



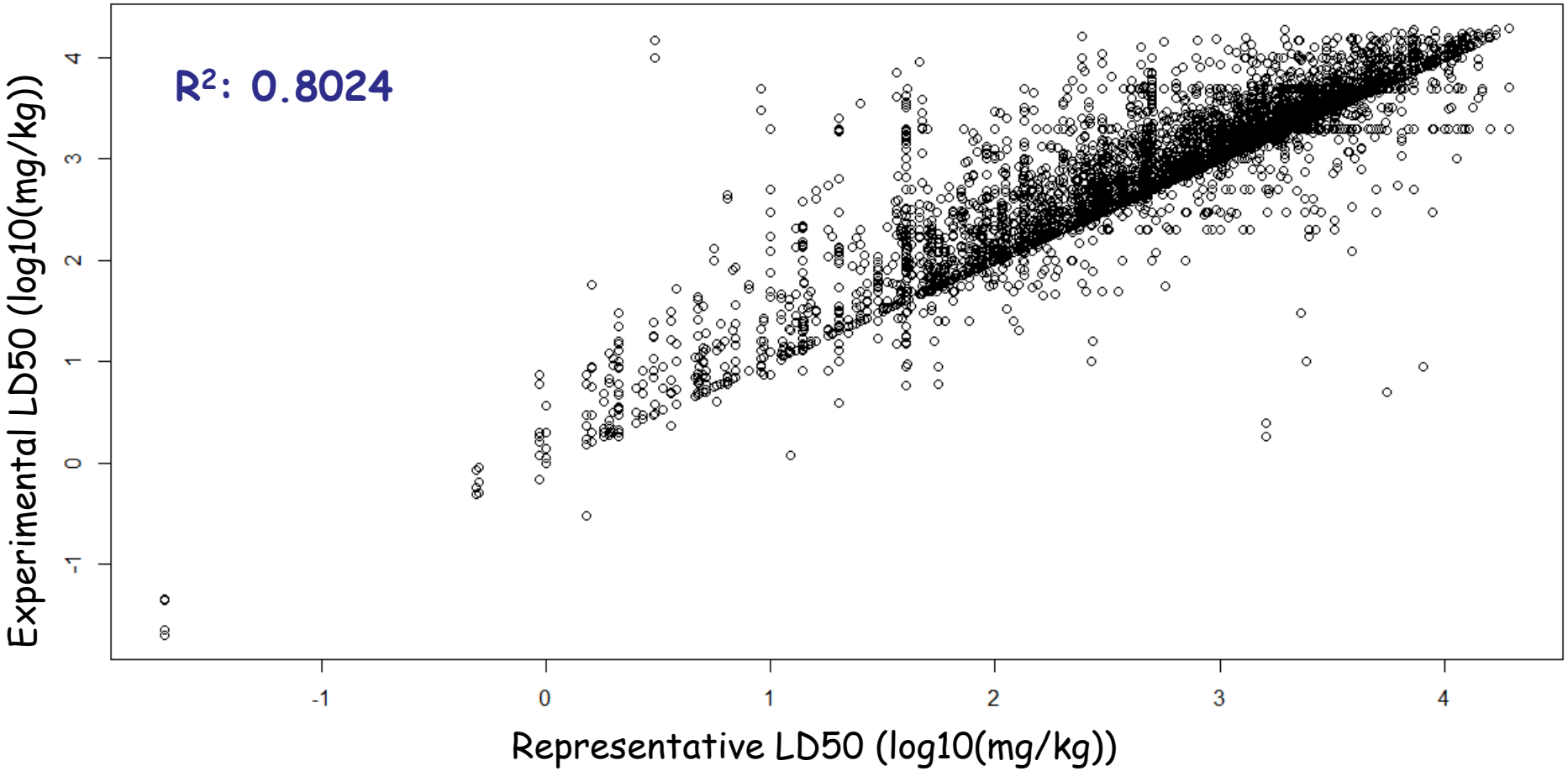
Defining a Confidence Range

Bootstrapping of the standard deviations for repeat test chemicals (~1120 with >3 replicates) identified a 95% confidence interval for LD50 values of ± 0.31 $\log_{10}$(mg/kg)



Assessing "Performance" of the Animal Assay

Representative LD50 vs. Experimental Values



RMSE of 0.42 was also computed for this dataset based on the LD50 values

Acute Toxicity Implementation Plan

- Coordinate activities via ICCVAM Workgroups
- Draft a scoping document to identify U.S. agency requirements, needs, and decision contexts for acute toxicity data
- Coordinate efforts with stakeholders
- Identify, acquire, and curate high quality data from reference test methods
- Identify and evaluate non-animal alternative approaches to acute toxicity testing
- Gain regulatory acceptance and facilitate use of non-animal approaches

Identify and evaluate non-animal alternative approaches to acute toxicity testing

- Establish a dataset of rat oral acute toxicity study LD50 data
- Evaluate the variability of the experimental data collected
  - to inform data curation efforts
  - to inform considerations for evaluating performance and coverage of existing models
  - to inform considerations for new model development
- Identify endpoints to be modeled based on US agency needs
- Evaluate existing models for acute toxicity
- Investigate the feasibility of developing new models for acute toxicity
- Initiate a project to leverage the expertise of the international modelling community to develop predictive models of acute oral toxicity
- Evaluate the applicability of the existing and new models for chemistries of interest to US agencies

Identify and evaluate non-animal alternative approaches to acute toxicity testing

- Establish a dataset of rat oral acute toxicity study LD50 data
- Evaluate the variability of the experimental data collected
  - to inform data curation efforts
  - to inform considerations for evaluating performance and coverage of existing models
  - to inform considerations for new model development
- Identify endpoints to be modeled based on US agency needs
- Evaluate existing models for acute toxicity
- Investigate the feasibility of developing new models for acute toxicity
- Initiate a project to leverage the expertise of the international modelling community to develop predictive models of acute oral toxicity
- Evaluate the applicability of the existing and new models for chemistries of interest to US agencies

Identify and evaluate non-animal alternative approaches to acute toxicity testing

- Evaluating existing *in silico* models

Model	Number of substances in dataset	Number of substances that could be predicted	Accuracy for substances with one Value	Accuracy for substances with multiple values	Overall Accuracy
TIMES Model	1787	315 (17.6%)	85 of 93 (91%)	206 of 222 (93%)	291 of 315 (92%)
TEST-Acute Oral Consensus Model	1787	1673 (93.6%)	433 of 490 (88%)	1092 of 1183 (92%)	1525 of 1673 (91%)

Fitzpatrick et al., Presented at ASCCT 2017; SOT 2018, manuscript in preparation

Identify and evaluate non-animal alternative approaches to acute toxicity testing

- Establish a dataset of rat oral acute toxicity study LD50 data
- Evaluate the variability of the experimental data collected
  - to inform data curation efforts
  - to inform considerations for evaluating performance and coverage of existing models
  - to inform considerations for new model development
- Identify endpoints to be modeled based on US agency needs
- Evaluate existing models for acute toxicity
- Investigate the feasibility of developing new models for acute toxicity
- Initiate a project to leverage the expertise of the international modelling community to develop predictive models of acute oral toxicity
- Evaluate the applicability of the existing and new models for chemistries of interest to US agencies

Identify and evaluate non-animal alternative approaches to acute toxicity testing

Informed by the Conceptual Framework outlined in the NRC 2015 report prepared for DOD

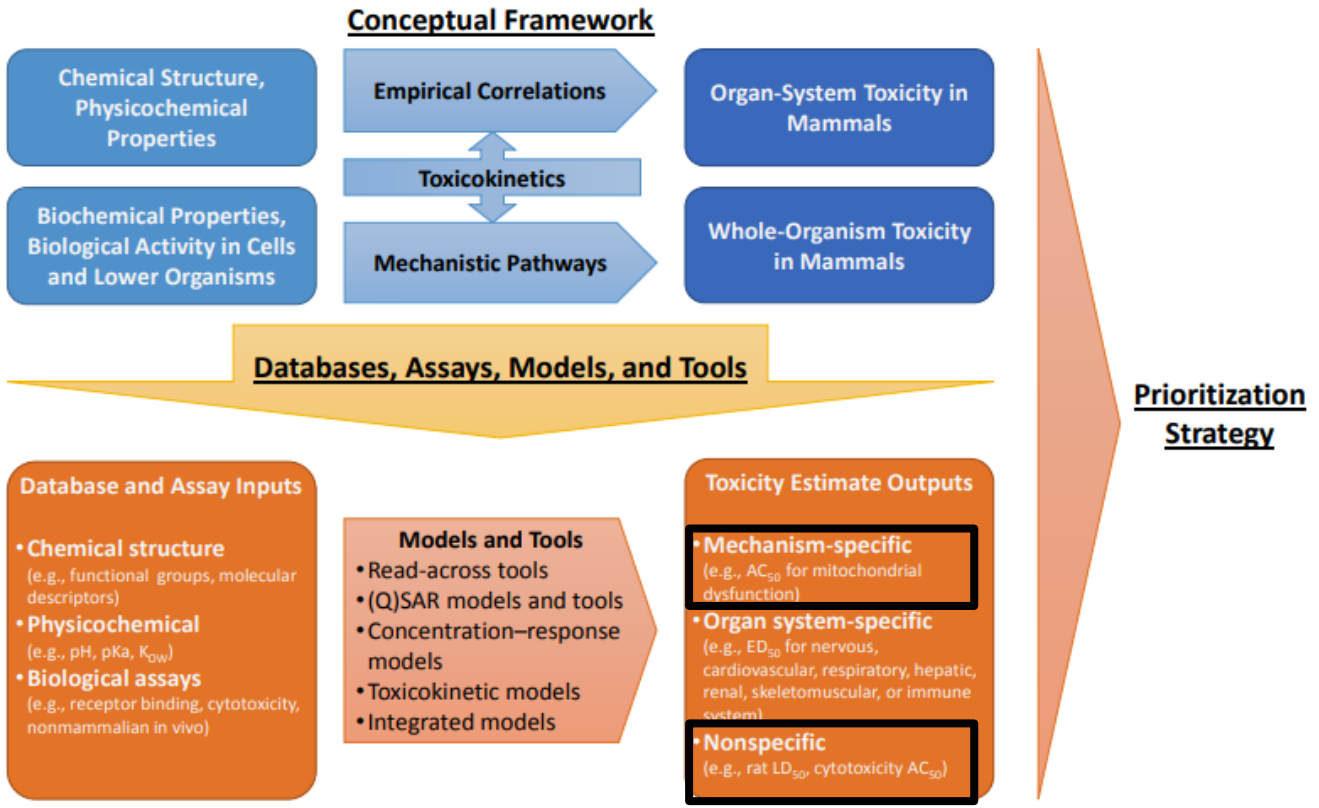


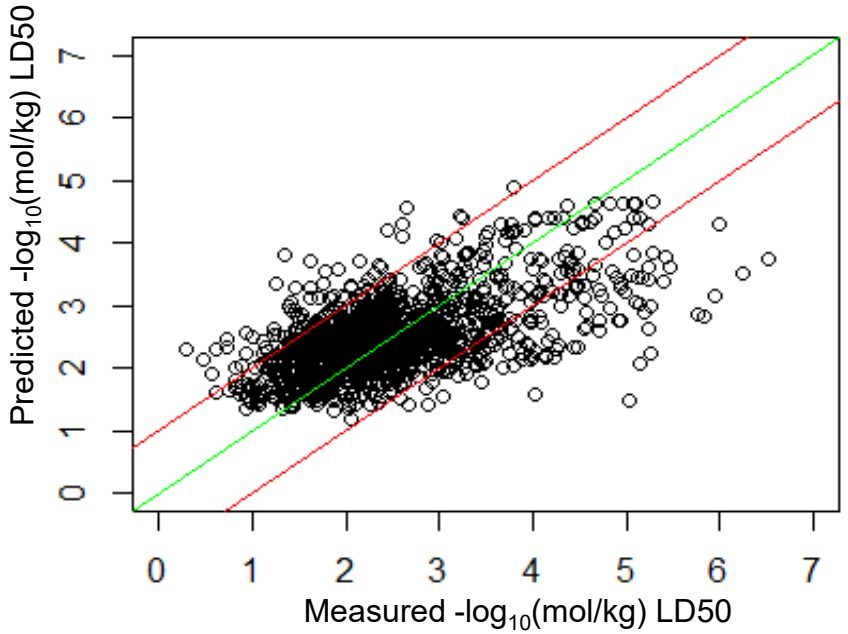
FIGURE S-1 Conceptual framework and examples of databases, assays, models, and tools for predicting acute chemical toxicity.

Identify and evaluate non-animal alternative approaches to acute toxicity testing

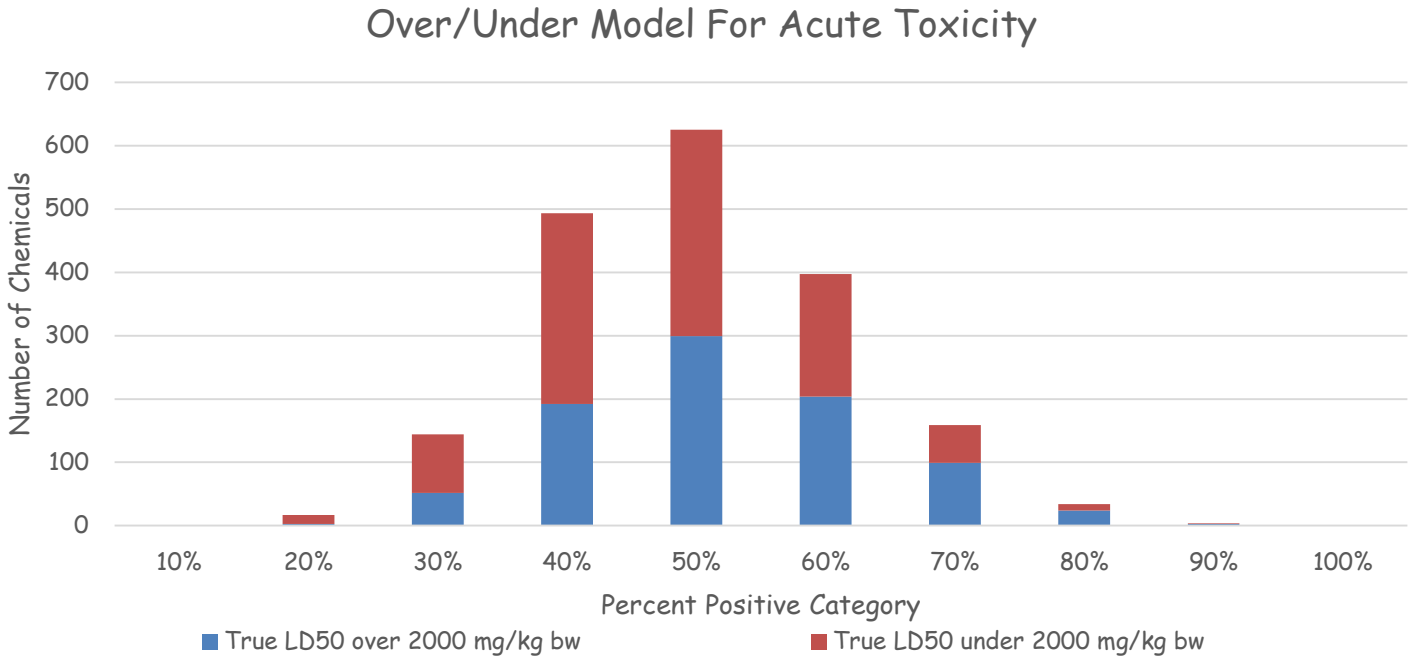
- NCCT efforts include:
- Developing new global models to predict LD50 or a toxic/non toxic category
- Local Cluster-based Regression Models based on chemical, biological, hybrid and MOA-chemical
- Read-across approaches using Generalised Read-across (GenRA)

Identify and evaluate non-animal alternative approaches to acute toxicity testing

- Developing new Global models:
- Global Regression Model



- Global Random Forest Model



- Global ridge regression model used both experimental and predicted ToxCast™ and Tox21 assay outcomes as descriptors.
- Training set (4164), Test set (1387)
- 85% of the substances were found to be within one log unit of their predicted LD50 value.

- Model for predicting compounds over and under a LD50 of 2000 mg/kg bw had an accuracy of 57%, a balanced accuracy of 56%, a sensitivity of 57%, and a specificity of 56%.

Fitzpatrick et al., Presented at ASCCT 2017; SOT 2018, manuscript in preparation

Identify and evaluate non-animal alternative approaches to acute toxicity testing

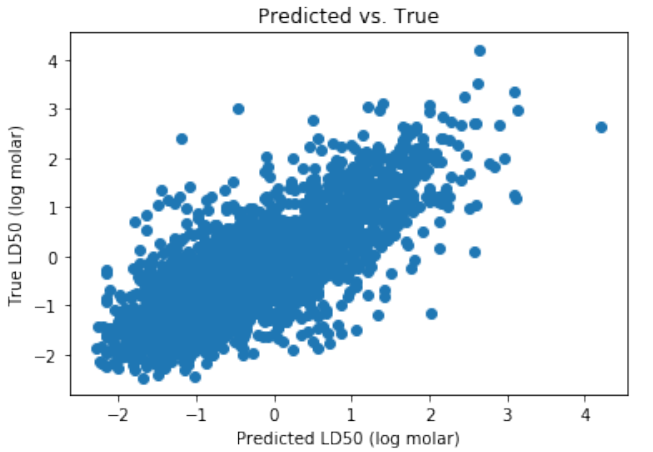
- **Developing new Local models:**
- **Local Cluster-based Regression Models – chemical, biological, hybrid and MOA-chemical**

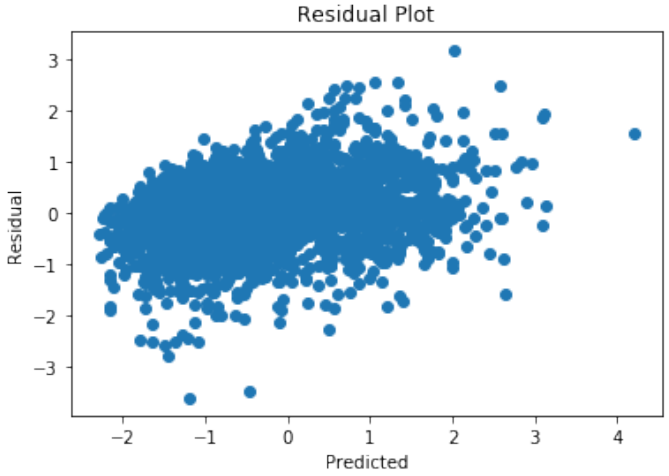
1. Chemical Clusters-Chemical QSAR	2. Chemical Clusters-Biological QSAR	3. Biological Clusters-Chemical QSAR	4. Chemical-Biological Clusters, QSAR	5. MOA Clusters-Chemical QSAR
• **Chemical Descriptors** – ToxPrints – PaDEL Descriptors – CDK Descriptors	• **Chemical Descriptors** – ToxPrints • **Biological Descriptors** – ToxCast Group B assays	• **Biological Descriptors** – ToxCast Group B assays • **Chemical Descriptors** – ToxPrints – PaDEL Descriptors – CDK Descriptors	• **Chemical Descriptors** – ToxPrints • **Biological Descriptors** – ToxCast Group B assays	• Replace each assay by MOA • Final MOA outcome: = 1, if chemical active in any assay for the given MOA threshold = 0, otherwise • **Chemical Descriptors** • ToxPrints • PaDEL Descriptors • CDK Descriptors

Pradeep et al., in preparation

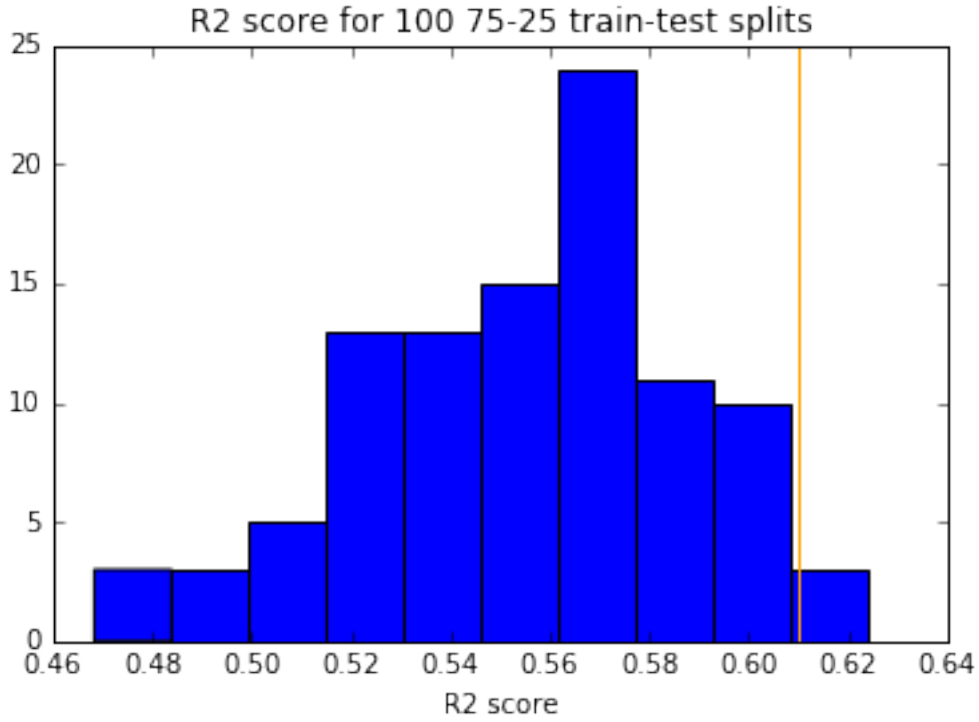
Identify and evaluate non-animal alternative approaches to acute toxicity testing

- Developing read-across models: using GenRA





- R^2 = 0.61
- RMSE = 0.58
- A few outliers, but not too extreme
- Residuals clustered around zero with no obvious patterns



- Estimate confidence in R2
- 75-25 train-test splits
- R^2 values range from 0.46 to 0.62

Helman et al., in press

Identify and evaluate non-animal alternative approaches to acute toxicity testing

- Establish a dataset of rat oral acute toxicity study LD50 data
- Evaluate the variability of the experimental data collected
  - to inform data curation efforts
  - to inform considerations for evaluating performance and coverage of existing models
  - to inform considerations for new model development
- Identify endpoints to be modeled based on US agency needs
- Evaluate existing models for acute toxicity
- Investigate the feasibility of developing new models for acute toxicity
- Initiate a project to leverage the expertise of the international modelling community to develop predictive models of acute oral toxicity
- Evaluate the applicability of the existing and new models for chemistries of interest to US agencies

Identify and evaluate non-animal alternative approaches to acute toxicity testing

- Use large database of rat oral LD50 values to train (and test) QSAR models to predict acute oral systemic toxicity
- 32 groups from the US, Europe, and Asia responded with 135 models for LD50, EPA and GHS categories, and binary nontoxic vs all others and very toxic vs all others.
- Models were qualitatively and quantitatively assessed and combined into consensus models.
- Consensus model performance compared with animal test reproducibility for binary, categorical, and quantitative models

https://ntp.niehs.nih.gov/go/tox-models

Agency-Based Modeling Endpoint Selection

Binary Models

Hazard

- Highly toxic (≤50 mg/kg)
- Toxic (>50-5000 mg/kg)

+ Nontoxic (>2000 mg/kg)

Continuous Model

Point estimates of LD50 values

Categorical Models

EPA Categories

Hazard

- I (≤ 50 mg/kg)
- II (>50 ≤ 500 mg/kg)
- III (>500 ≤ 5000 mg/kg)
- IV (>5000 mg/kg)

GHS Categories

Packing Group

- I (≤ 5 mg/kg)
- II (>5 ≤ 50 mg/kg)
- III (>50 ≤ 300 mg/kg)
- IV (>300 ≤ 2000 mg/kg)
- NC (> 2000 mg/kg)

OSHA Hazard

Slide from K Mansouri

Performance Assessment

Consensus Model Statistics

	Very Toxic		Non-Toxic		EPA		GHS	
	Train	Eval	Train	Eval	Train	Eval	Train	Eval
Sensitivity	0.87	0.67	0.93	0.70	0.73	0.50	0.63	0.45
Specificity	0.94	0.96	0.96	0.88	0.96	0.91	0.91	0.92
Balanced Accuracy	0.93	0.81	0.94	0.79	0.83	0.71	0.77	0.68
***In vivo* Balanced Accuracy**	0.81		0.89		0.82		0.79	

	LD50 values		LD50 values
	Train	Eval	*In Vivo*
R2	0.84	0.64	0.80
RMSE	0.32	0.51	0.42

The consensus predictions perform just as well as replicate *in vivo* data do at predicting oral acute toxicity outcome

Slide from K Mansouri

Insights from the workshop

- Consensus model was equivalent in performance to the ability of the rat oral data LD50 to predict itself
- Is the reproducibi erformance threshold for alte
- Is that sufficient
- Are there other w etter communicated to s
- Desire for mechar 50 when the test itself does not pr nts in chemical structure associations with toxicity would be helpful
- Ways in which modelling approaches and outcomes can be captured in clearly defined workflows to facilitate communication and interpretation
- Demonstrating model performance for reference lists (inventories) pertinent to specific regulatory programmes to highlight relevance and utility

Computational Toxicology 8 (2018) 21–24

Contents lists available at ScienceDirect

Computational Toxicology

journal homepage: www.elsevier.com/locate/comtox

Predictive models for acute oral systemic toxicity: A workshop to bridge the gap from research to regulation

Nicole C. Kleinstreuer[a], Agnes L. Karmaus[b], Kamel Mansouri[b], David G. Allen[b], Jeremy M. Fitzpatrick[c], Grace Patlewicz[c,*]

Summary remarks

- Outlined ATWG charges
- Substantial progress has been made in outlining the decision contexts, needs and gathering the acute toxicity data to inform the array of in silico modelling efforts
- Evaluating the variability of the acute toxicity data is a key consideration both in terms of the impact this has in current hazard assessments but also in managing expectations of the performance of new models
- Various models have been developed which have reasonable performance but a consensus approach provides a way of taking the 'best' from many models [see next presentation]
- Next steps for the ATWG include gathering data on acute inhalation data, evaluating the data and investigating the feasibility of developing new models