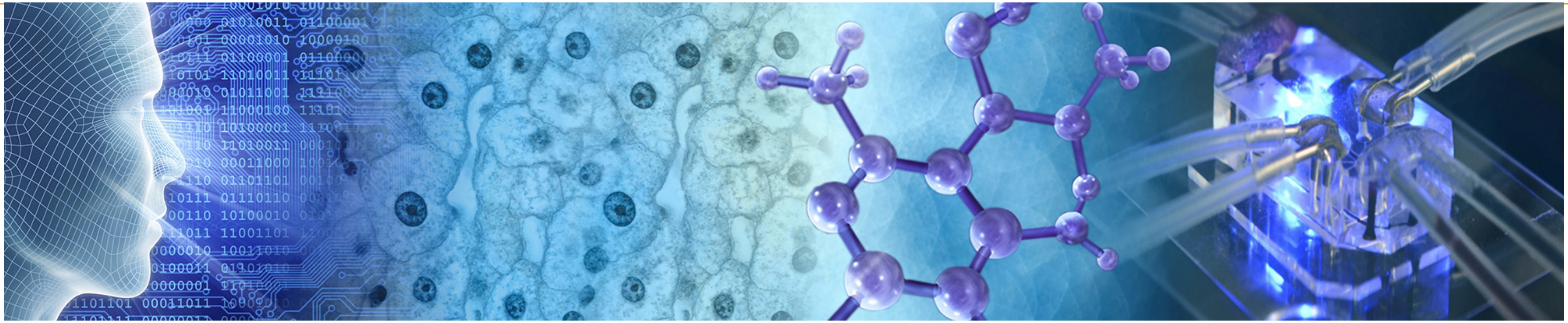
NTP
National Toxicology Program



Collaborative Acute Toxicity Modeling Suite (CATMoS)

Kamel Mansouri

LEAD COMPUTATIONAL CHEMIST, ILS IN SUPPORT OF NICEATM

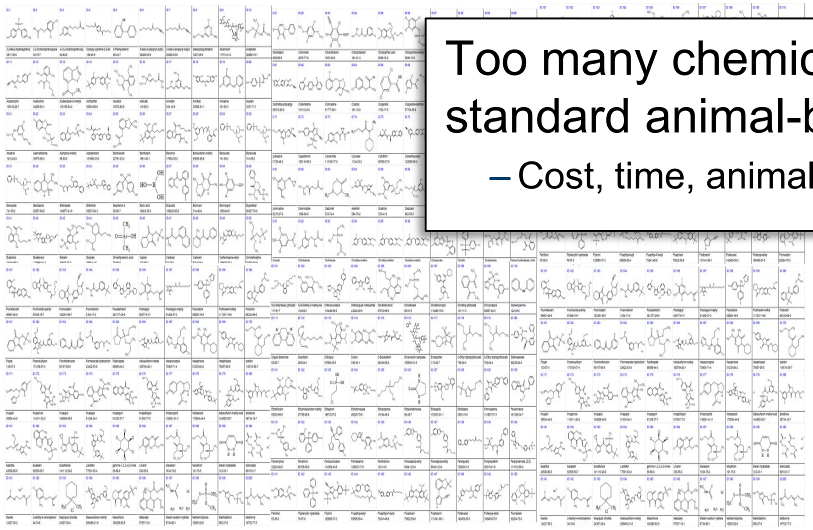
SACATM – September 19-20, 2019

Disclaimer: ILS staff provide technical support for NICEATM,
but do not represent NIEHS, NTP, or the official positions of any federal agency.
(the author declares no conflict of interest)

Overview

- Project scope: acute oral toxicity
  - Regulatory use of these data
  - Endpoints selected for predictive modeling
  - Compiling inventory of rat acute oral LD50
  - Establishing training, evaluation, and prediction sets
  - Evaluation of submitted models
- International contributors
- Generation of consensus predictions
- Current status and public release

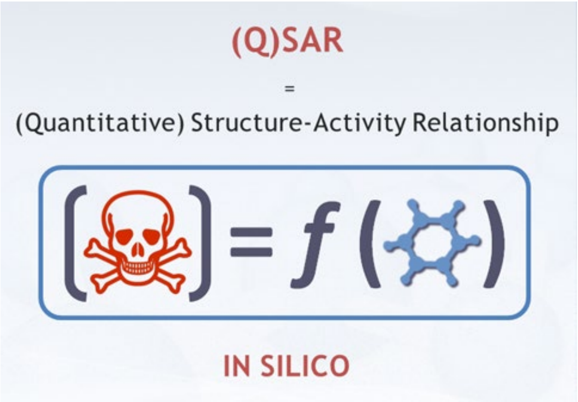
Toxicity prediction



Too many chemicals to test with standard animal-based methods

– Cost, time, animal welfare

Alternative



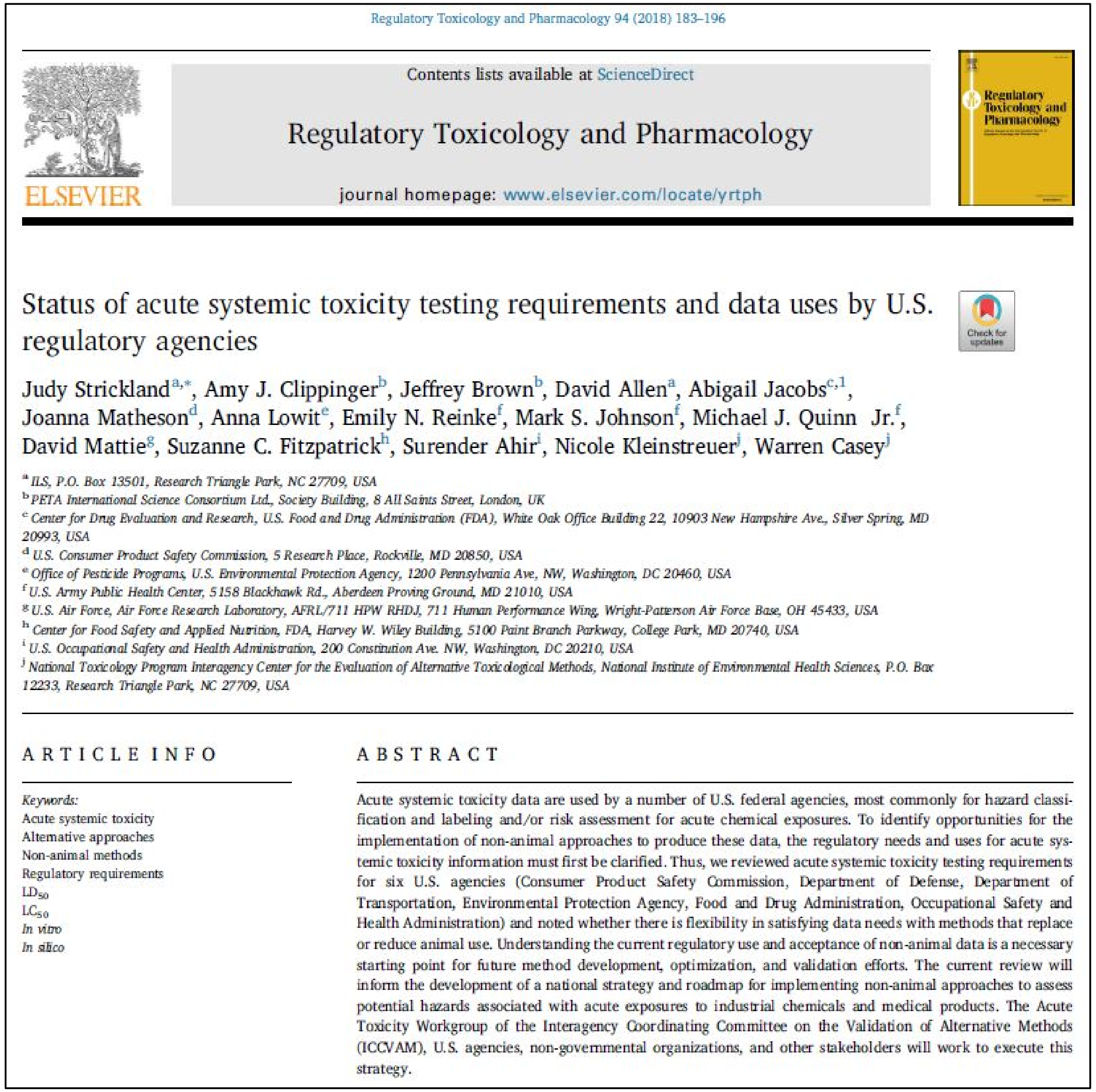
- Organic **pollutants** with exposure potential **accumulate** in body tissues
  - ➢ Cause **toxic effects** to wild life and humans
- Existence of **gaps in the experimental data** for environmental endpoints
  - ➢ Need to fill the data gaps and bridge the **lack of knowledge**
- **Regulatory** requirements:
  - ➢ Reduce **animal** testing**, time** and **costs**

➢ **Methodology:** use of **QSAR/QSPR** to **predict** the **endpoints** of interest.

Scoping Regulatory Needs

ICCVAM Acute Toxicity Workgroup

- Identifies federal agency requirements, needs, and decision contexts for using acute systemic toxicity data



Regulatory Toxicology and Pharmacology 94 (2018) 183–196

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

ELSEVIER

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]

[a] ILS, P.O. Box 13501, Research Triangle Park, NC 27709, USA
[b] PETA International Science Consortium Ltd., Society Building, 8 All Saints Street, London, UK
[c] Center for Drug Evaluation and Research, U.S. Food and Drug Administration (FDA), White Oak Office Building 22, 10903 New Hampshire Ave., Silver Spring, MD 20993, USA
[d] U.S. Consumer Product Safety Commission, 5 Research Place, Rockville, MD 20850, USA
[e] Office of Pesticide Programs, U.S. Environmental Protection Agency, 1200 Pennsylvania Ave, NW, Washington, DC 20460, USA
[f] U.S. Army Public Health Center, 5158 Blackhawk Rd., Aberdeen Proving Ground, MD 21010, USA
[g] U.S. Air Force, Air Force Research Laboratory, AFRL/711 HPW RHDJ, 711 Human Performance Wing, Wright-Patterson Air Force Base, OH 45433, USA
[h] Center for Food Safety and Applied Nutrition, FDA, Harvey W. Wiley Building, 5100 Paint Branch Parkway, College Park, MD 20740, USA
[i] U.S. Occupational Safety and Health Administration, 200 Constitution Ave. NW, Washington, DC 20210, USA
[j] National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709, USA

ARTICLE INFO

Keywords:
Acute systemic toxicity
Alternative approaches
Non-animal methods
Regulatory requirements
LD_{50}
LC_{50}
In vitro
In silico

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.

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

Hazard

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

Available data for modeling

Rat oral LD50s:
16,297 chemicals total
34,508 LD50 values

15,688 chemicals total
21,200 LD50 values

QSAR-ready standardization →

Desalted, stereochemistry stripped, tautomers and nitro groups standardized, valence corrected, structures neutralized

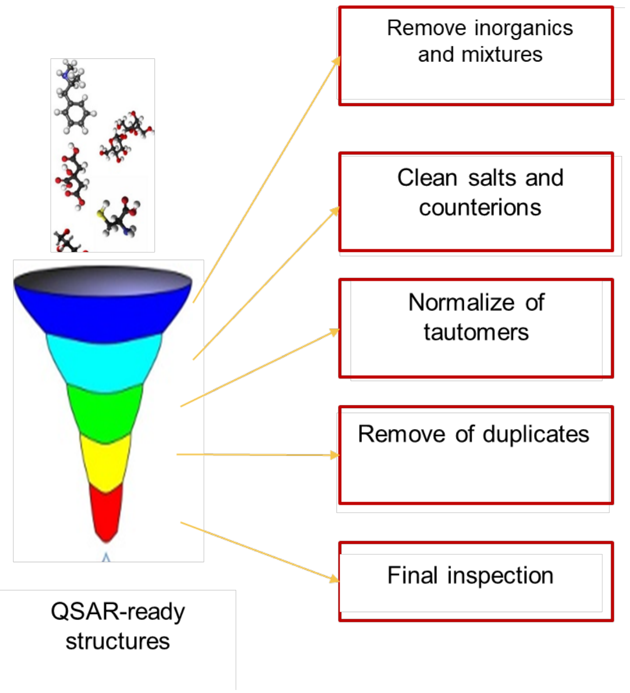
11992 chemicals with accurate structures

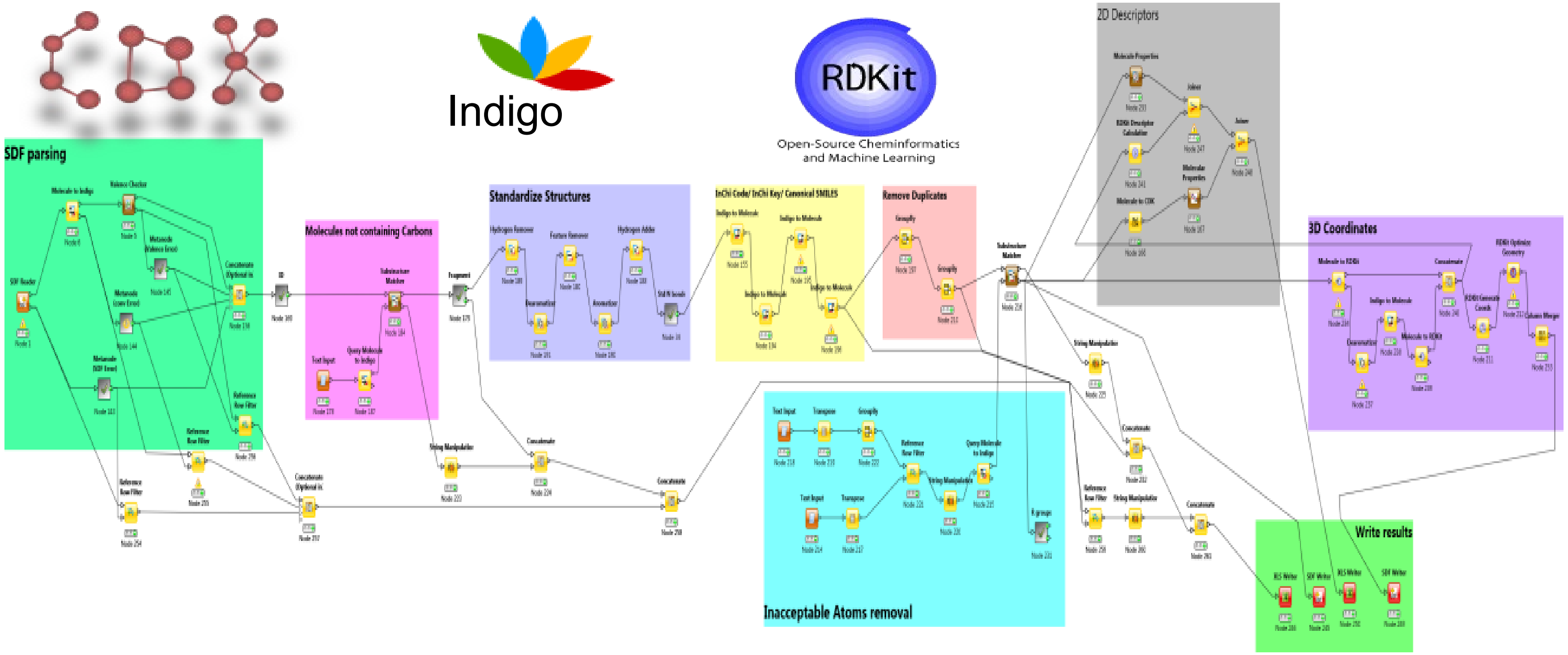
- Very toxic endpoint: 11886 entries (binary, 0/1)
- Non-toxic endpoint: 11871 entries (binary, 0/1)
- EPA endpoint: 11755 entries (categorical, 4 categories)
- GHS endpoint: 11845 entries (categorical, 5 categories)
- LD50 endpoint: 8908 entries (continuous values)

QSAR-ready KNIME workflow

Aim of the workflow:

- Combine different procedures and ideas
- Minimize the differences between the structures used for prediction
- Produce a flexible free and open source workflow to be shared





Fourches et al. J Chem Inf Model, 2010, 29, 476 – 488

Wedebye et al. Danish EPA Environmental Project No. 1503, 2013

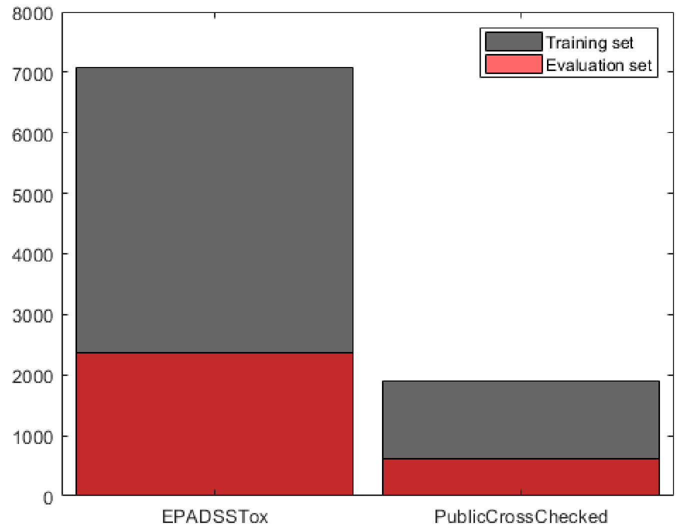
Mansouri et al. (http://ehp.niehs.nih.gov/15-10267/)

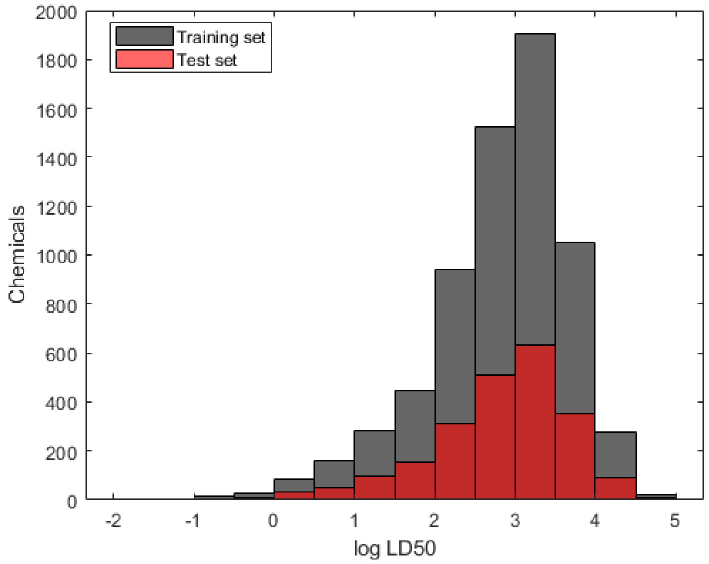
Establishing Modeling Dataset

- **Training and evaluation sets:**

- 11,992 chemicals from the final inventory of chemicals with QSAR-ready structures having rat oral acute toxicity data were split into training and test sets:
  - 75% training set: 8,994 chemicals
  - 25% evaluation set: 2,998 chemicals

- All endpoints training data included in same structure file
- Similar distributions and variability for values and categories
- Similar distribution of chemical structures sources





Establishing Modeling Dataset

- **Prediction set:**

Included lists of regulatory interest:

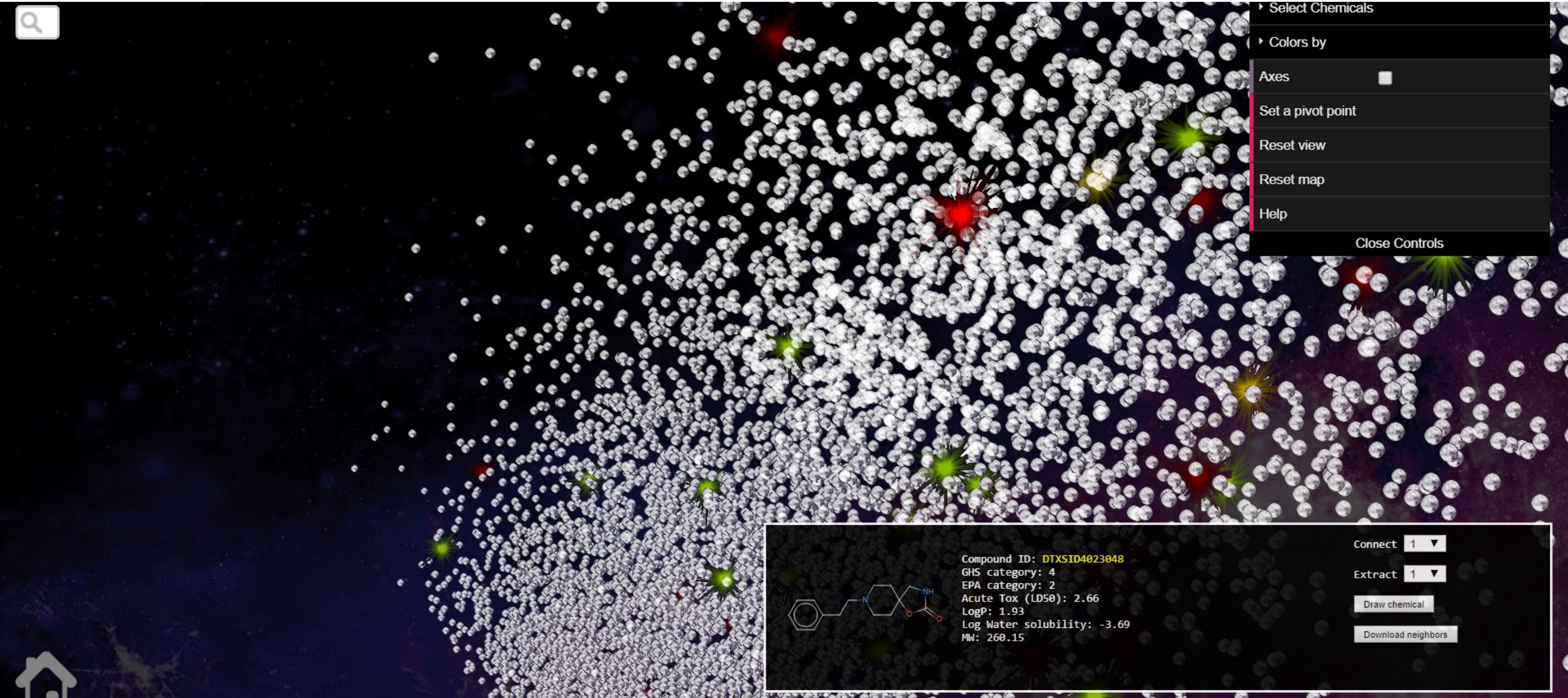
- ToxCast/Tox21
- EDSP
- TSCA
- Substances on the market (EPA Dashboard list)

After QSAR-ready standardization:

48137 structures to be predicted (including the evaluation set)

ChemMaps landscape of CATMoS chemicals

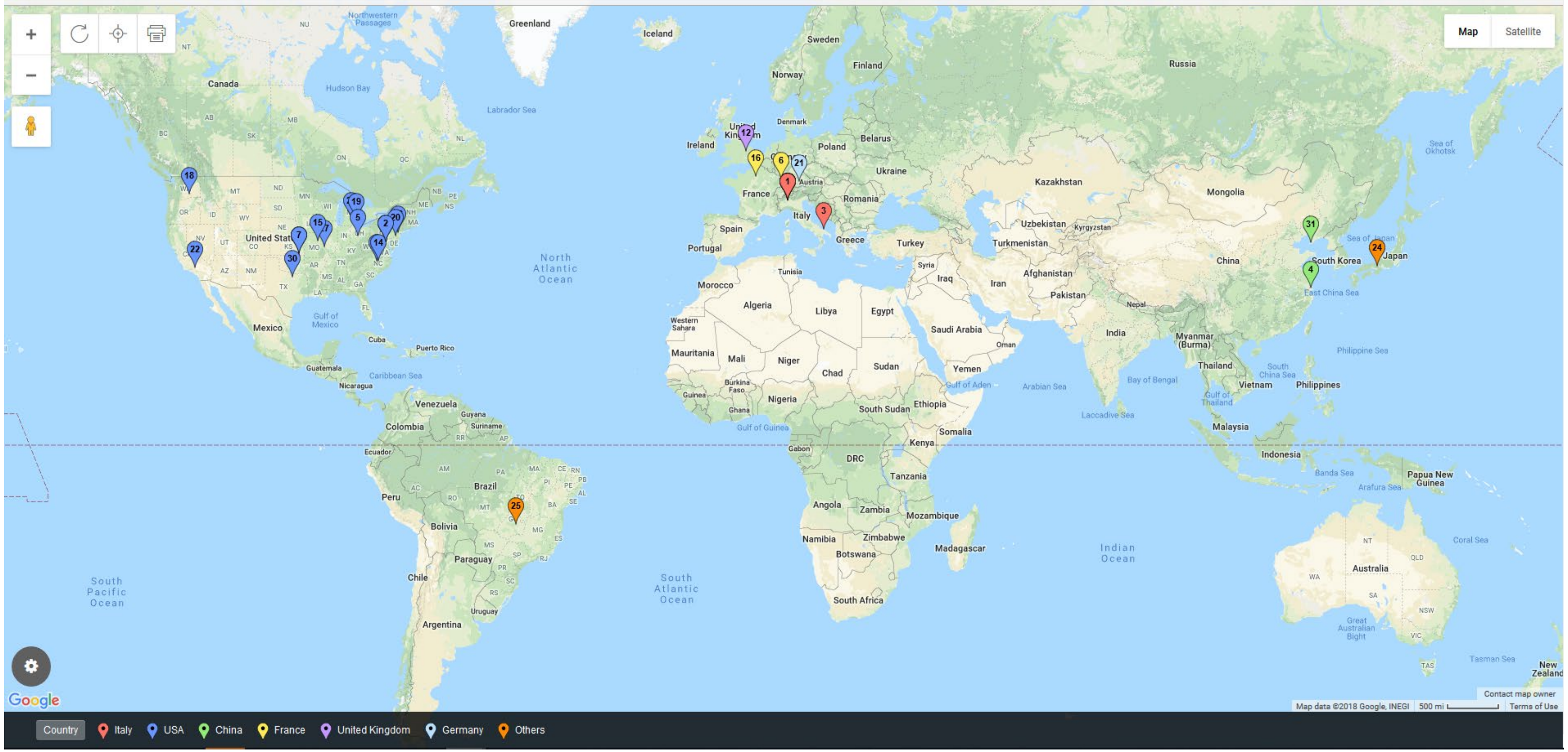
http://www.chemmaps.com/chemmaps/DSSToxMap3D/



International Collaboration

Consortium:

- **<u>35 Participants/Groups</u>** from around the globe representing academia, industry, and government contributed

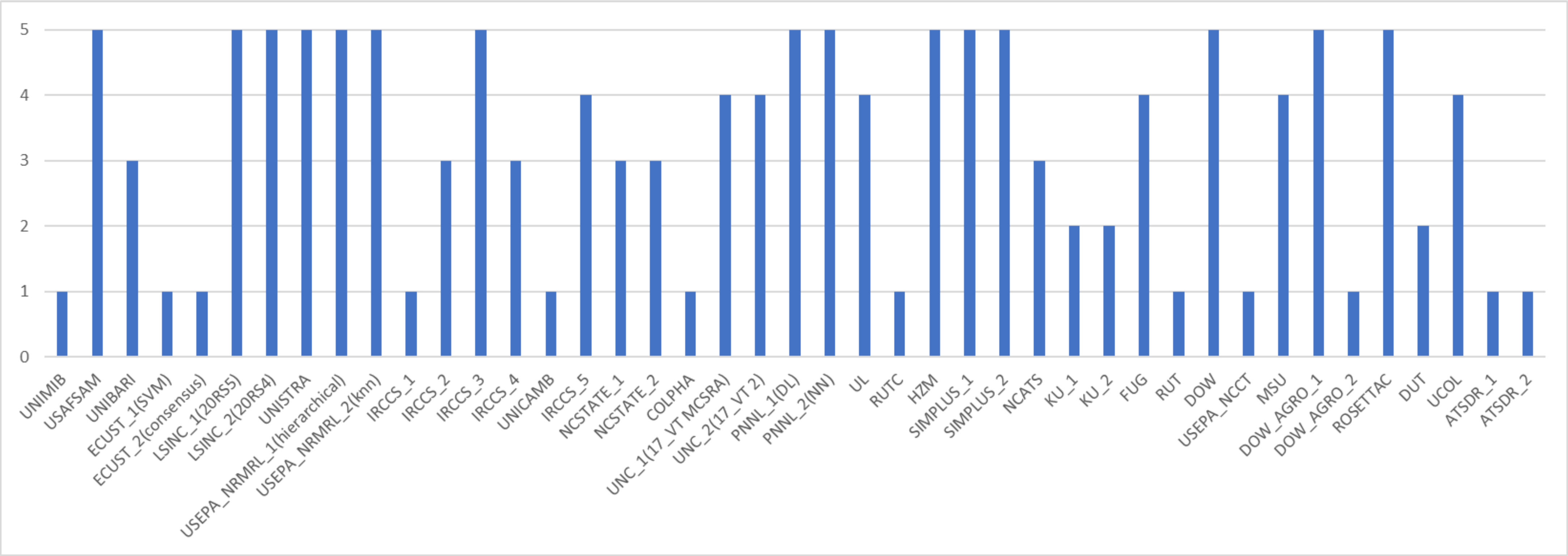


(https://batchgeo.com/map/d06c5d497ed8f76ecfee500c2b0e1dfa)

Submitted Models

- Non-toxic: 33 models
- Very Toxic: 32 models
- GHS categories: 23 models
- EPA categories: 26 models
- LD50: 25 models

Total: 139 models



Evaluation procedure

Qualitative evaluation:

- Documentation
- Defined endpoint
- Unambiguous algorithm
- Availability of code
- Applicability domain definition
- Availability of data used for modeling
- Mechanistic interpretation

Quantitative evaluation:

- Goodness of fit: training statistics
- Evaluation set predictivity: statistics on the evaluation set
- Robustness: balance between (Goodness of fit) & (Test set predictivity)

$$S = 0.3 * (Goodness\ of\ fit) + 0.45 * (Test\ set\ predictivity) + 0.25 * (Robustness)$$

Categorical models (binary and multi-class):

$$Goodness\ of\ fit = 0.7 * (BA_{Tr}) + 0.3 * \left(1 - |Sn_{Tr} \widetilde{-} Sp_{Tr}|\right)$$

$$Test\ set\ predictivity = 0.7 * (BA_{Tst}) + 0.3 * \left(1 - |Sn_{Tst} \widetilde{-} Sp_{Tst}|\right)$$

$$Robustness = 1 - |BA_{Tr} - BA_{Tst}|$$

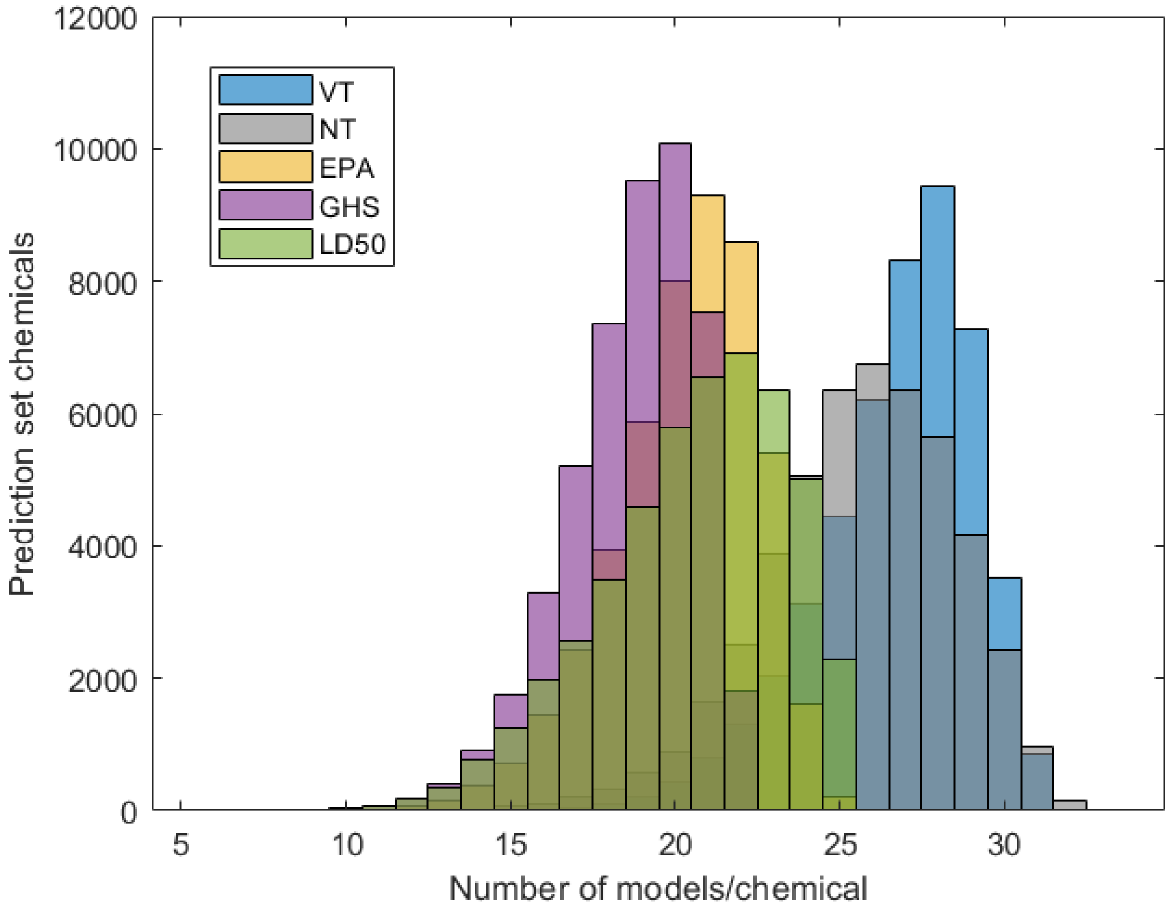
Continuous models:

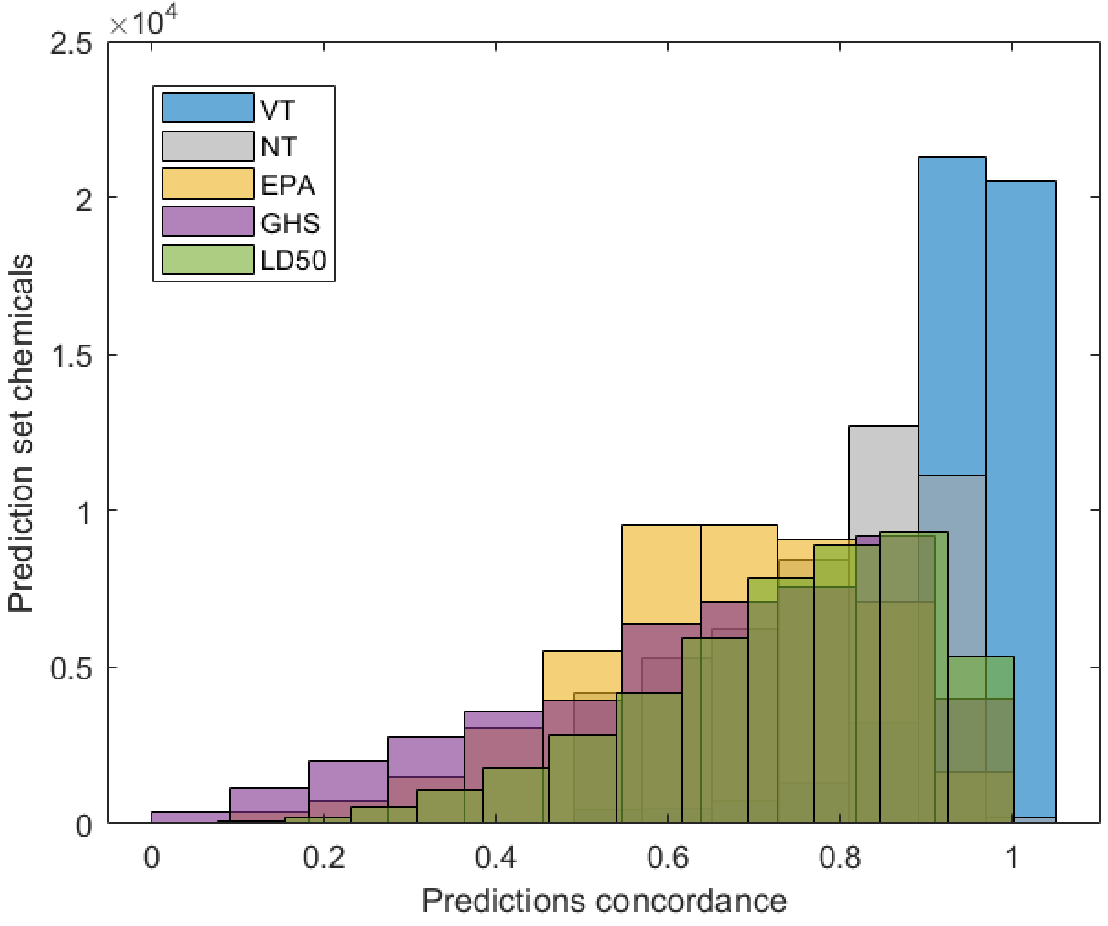
$$Goodness\ of\ fit = R^2_{Tr}$$

$$Test\ set\ predictivity = R^2_{Tst}$$

$$Robustness = 1 - |R^2_{Tr} - R^2_{Tst}|$$

Coverage and concordance of the models





CATMoS consensus modeling

Steps of combining the single models into consensus

Initial models & predictions

- VT (32 models)
- NT (33 models)
- GHS (23 models)
- EPA (26 models)
- LD50 (25 models)

Combining models

Step 1

Weighted average /majority rule

Independent consensus models/predictions

- VT
- NT
- GHS
- EPA
- LD50

A consensus model per endpoint (~20-~30 models)

Weight of Evidence approach (WoE)

Step 2

Majority rule

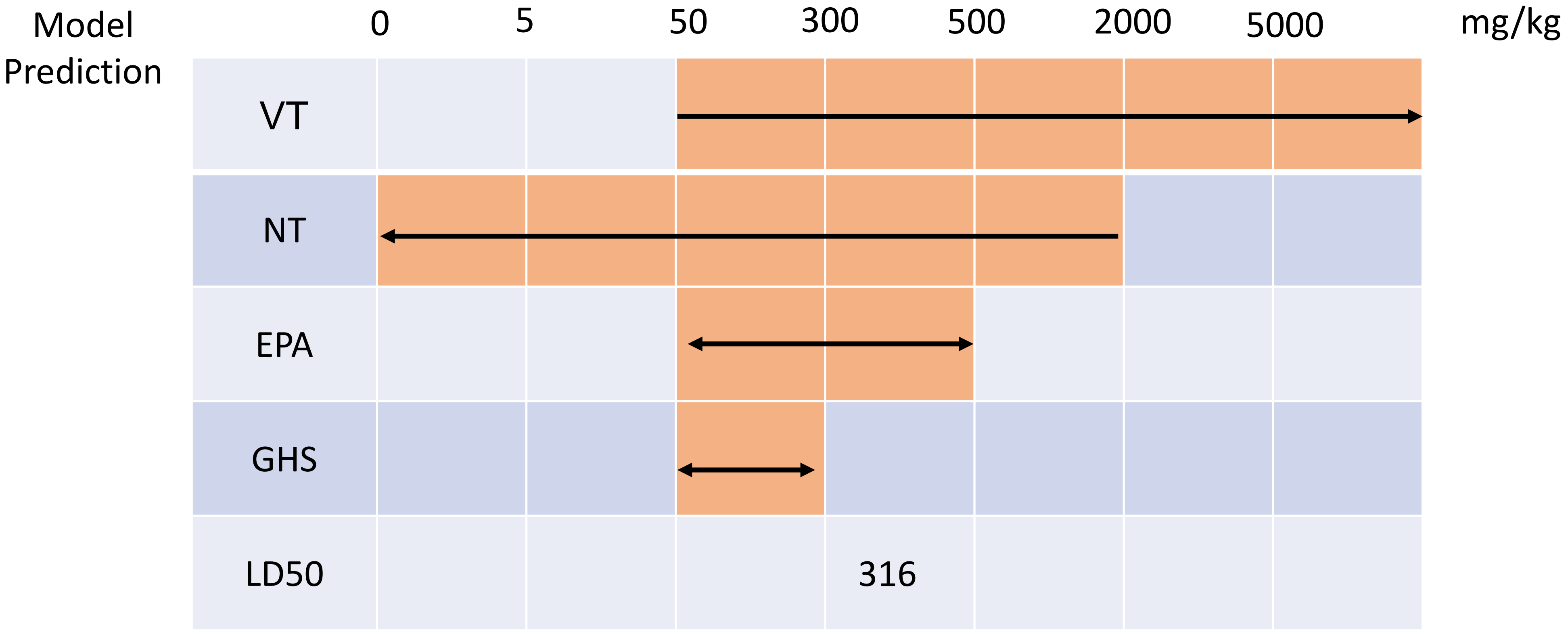
Consistent consensus models/predictions

- VT
- NT
- GHS
- EPA
- LD50

Consensus representing all ~140 models

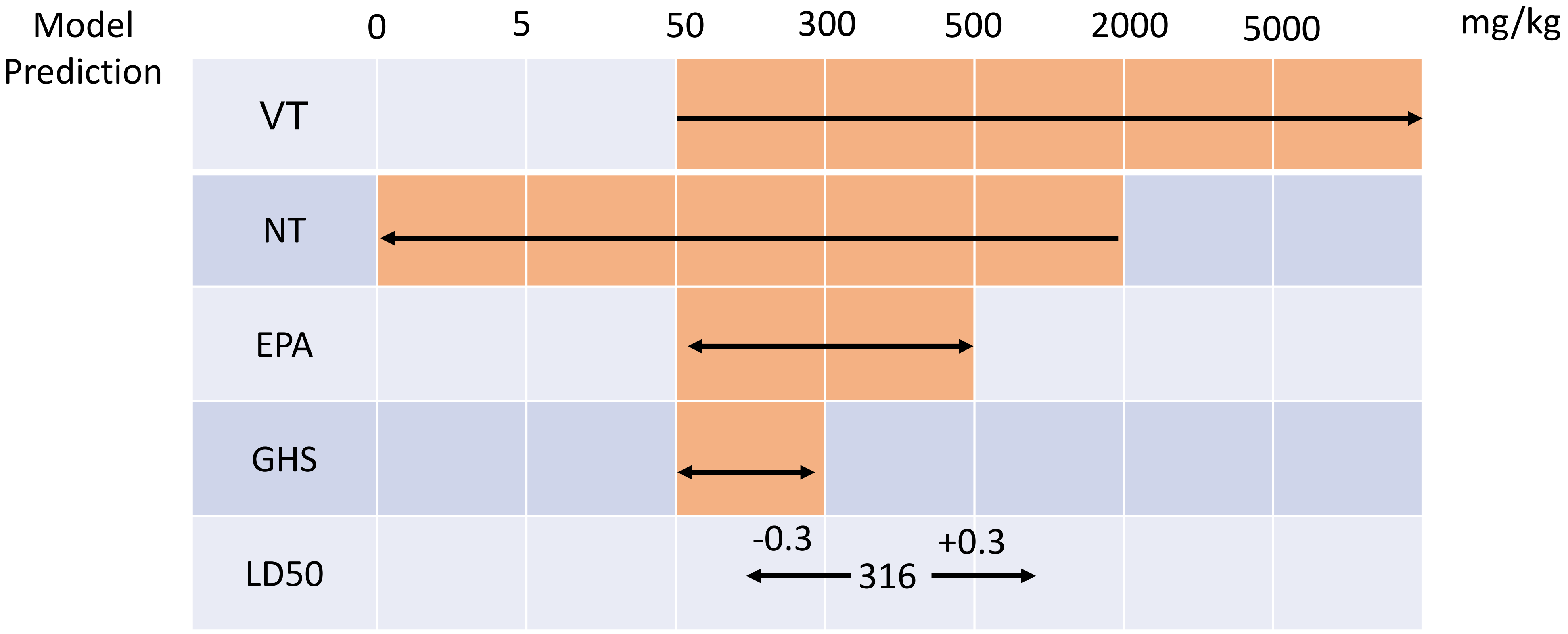
WoE approach to combine the 5 endpoints

	VT	NT	EPA	GHS	LD50
molX	0	0	2	3	2.5



WoE approach to combine the 5 endpoints

	VT	NT	EPA	GHS	LD50
molX	0	0	2	3	2.5



Variability range (log units) for LD50

WoE approach to combine the 5 endpoints

	VT	NT	EPA	GHS	LD50
molX	0	0	2	3	2.5

Model Prediction	0	5	50	300	500	2000	5000 mg/kg
VT	0	0	1	1	1	1	1
NT	1	1	1	1	1	0	0
EPA	0	0	1	1		0	0
GHS	0	0	1	0	0	0	0
LD50	0	0	1	1	1		
WoE	**1**	**1**	**5**	**4**	**3**	**1**	**1**

WoE approach to combine the 5 endpoints

Original: independent calls

	VT	NT	EPA	GHS	LD50
molX	0	0	2	3	2.5

WoE: consistent calls

	VT	NT	EPA	GHS	LD50
molX	0	0	2	3	2.36

Winning bin

Model Prediction — 0, 5, 50, 300, 500, 2000, 5000 mg/kg

Model Prediction	0–5	5–50	50–300	300–500	500–2000	2000–5000	5000+
VT	0	0	1	1	1	1	1
NT	1	1	1	1	1	0	0
EPA	0	0	1	1		0	0
GHS	0	0	1	0	0	0	0
LD50?	0	0	1 160 ⟷	1	1 613		
WoE	**1**	**1**	**5**	**4**	**3**	**1**	**1**

How to adjust quantitative LD50?
Avg of Lower CI and upper bin threshold

(160+300)/2
=230mg/kg

Performance Assessment

Consensus Model Statistics

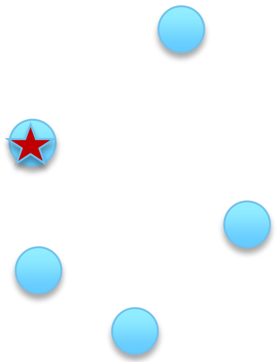
	VT Train	VT Eval	NT Train	NT Eval	EPA Train	EPA Eval	GHS Train	GHS Eval
Sensitivity	0.87	0.70	0.88	0.67	0.81	0.62	0.80	0.58
Specificity	0.99	0.97	0.97	0.90	0.92	0.86	0.95	0.90
Balanced Accuracy	0.93	0.84	0.92	0.78	0.87	0.74	0.88	0.74
***In vivo* Balanced Accuracy**	0.81		0.89		0.82		0.79	

	LD50 Train	LD50 Eval	LD50 *In Vivo*
R2	0.85	0.65	0.80
RMSE	0.30	0.49	0.42

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

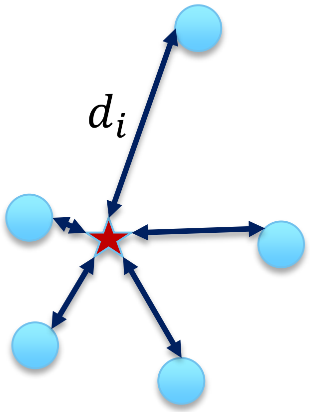
Extended CATMoS predictions

Weighted read-across



$$d_1 = 0$$

$$Pred_i = N_i$$



$$d_1 \neq 0$$

$$w_i = f(d_i)$$

$$Pred_i = f(w_i, N_i)$$

★ New chemical to be predicted

● Nearest neighbors (N_i)

d_i: Euclidean distance based on the selected descriptors for each endpoint

➡ Automated, similarity-endpoint dependent read-across: weighted kNN

Generation of Consensus Predictions

- Models passing qualitative evaluation (requirement for transparency; description of approach was sufficient)
- Integrating only *in-domain* predictions across chemicals in the prediction set (48,137 chemicals) for each model, respectively
  - Categorical models: weighted majority rule
  - Continuous model: weighted average

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,*]

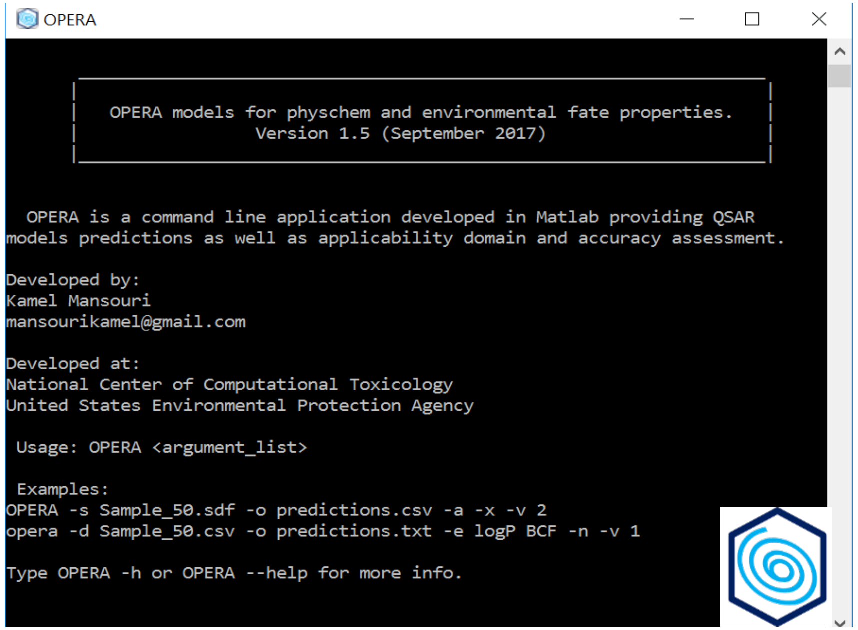
Collaboration with ATWG partners and ICCVAM agencies

Agency	No. Substances	Agency	No. Substances
Air Force	421	EPA OPP	36
Army Public Health Command	18	EPA OPPT	8
Army Edgewood Chemical Biological Center	42	EPA NCCT	4815
CPSC	110	FDA CFSAN	22
DOT	3671		

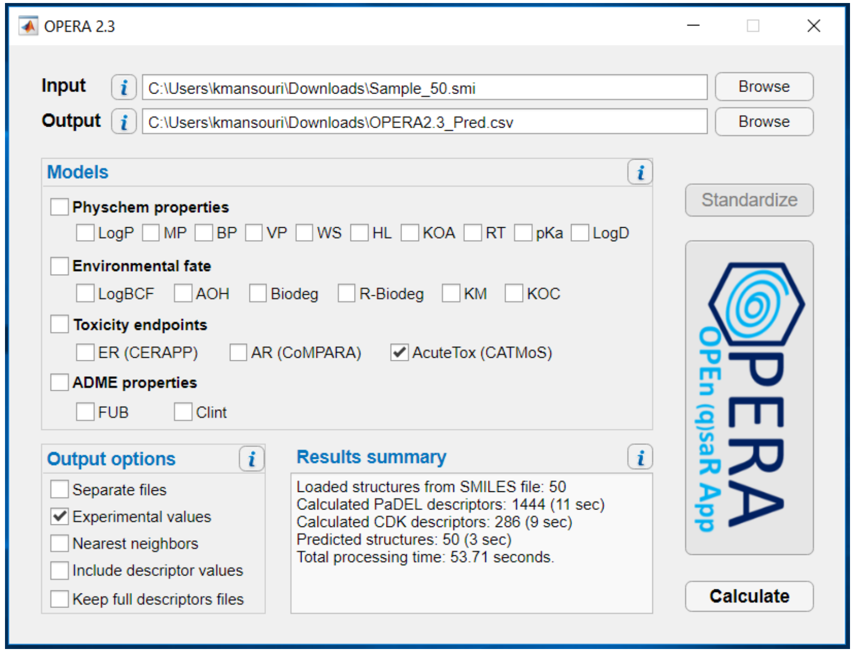
Evaluate and optimize CATMoS predictions based on lists of interest

Running CATMoS Consensus models

OPERA Standalone application



Command line



Graphical user interface

- Free, opensource & open-data
- Single chemical and batch mode
- Multiple platforms (Windows and Linux)
- Embeddable libraries (java, C, C++, Python)

https://github.com/NIEHS/OPERA

Mansouri et al. J Cheminform (2018). https://doi.org/10.1186/s13321-018-0263-1

OPERA2

OPERA 1.5

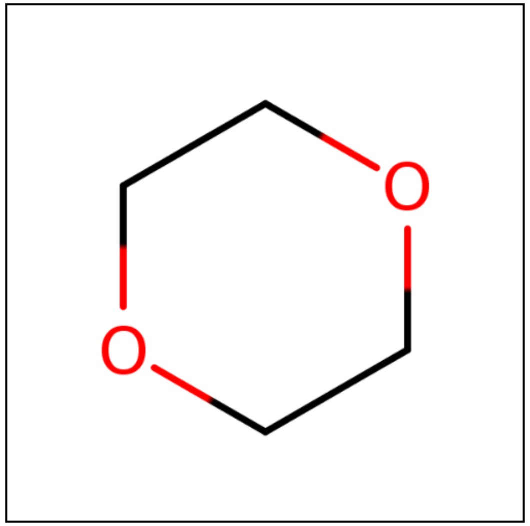
Physchem & Environmental fate:

Model	Property
AOH	Atmospheric Hydroxylation Rate
BCF	Bioconcentration Factor
BioHL	Biodegradation Half-life
RB	Ready Biodegradability
BP	Boiling Point
HL	Henry's Law Constant
KM	Fish Biotransformation Half-life
KOA	Octanol/Air Partition Coefficient
LogP	Octanol-water Partition Coefficient
MP	Melting Point
KOC	Soil Adsorption Coefficient
VP	Vapor Pressure
WS	Water solubility
RT	HPLC retention time

New in OPERA2:

- Physchem properties:
  - General structural properties
  - pKa
  - Log D
- ADME properties
  - Plasma fraction unbound (FuB)
  - Intrinsic clearance (Clint)
- Toxicity endpoints
  - ER activity (CERAPP) https://ehp.niehs.nih.gov/15-10267/
  - AR activity (CoMPARA) https://doi.org/10.13140/RG.2.2.19612.80009
  - **Acute toxicity (CATMoS) https://doi.org/10.1016/j.comtox.2018.08.002)**

CATMoS prediction examples



1,4-Dioxane

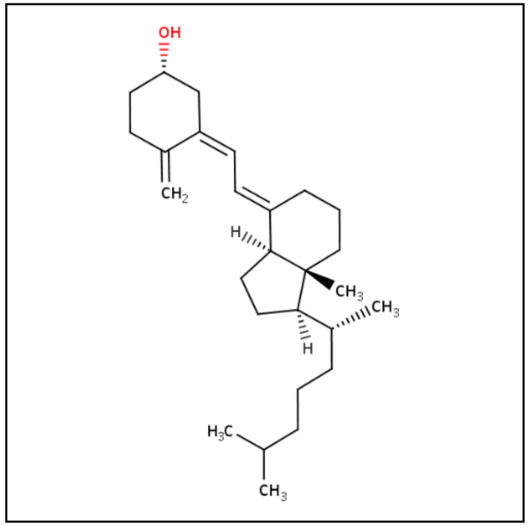
123-91-1 | DTXSID4020533

Molecular Formula: $C_4H_8O_2$

Average Mass: 88.106 g/mol

LD50: 4200 mg/kg
log10 LD50= 3.62

https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID4020533



Vitamin D3

67-97-0 | DTXSID6026294

Molecular Formula: $C_{27}H_{44}O$

Average Mass: 384.648 g/mol

LD50: 42 mg/kg
log10 LD50= 1.62

https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID6026294

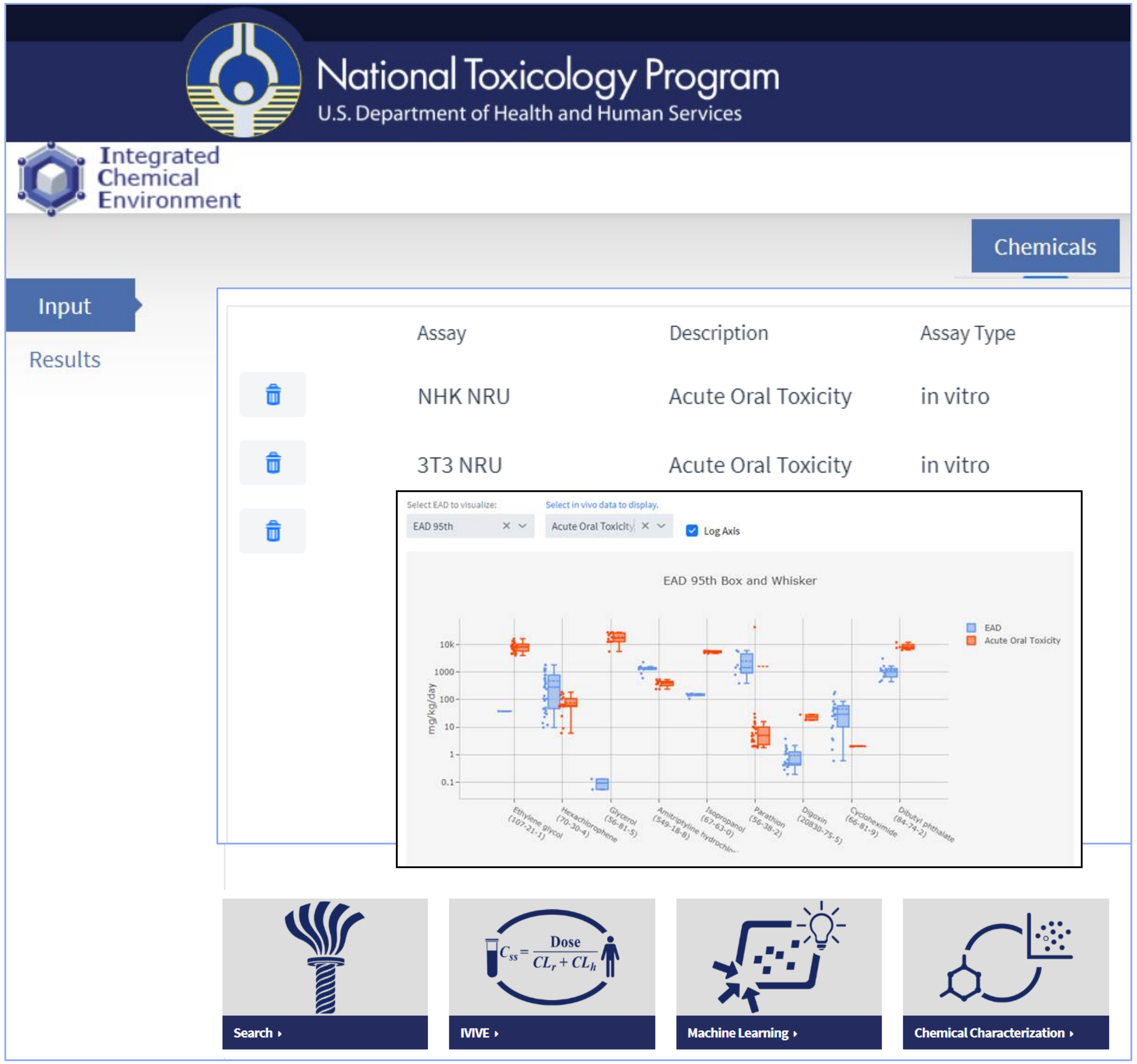
CATMoS predictions:

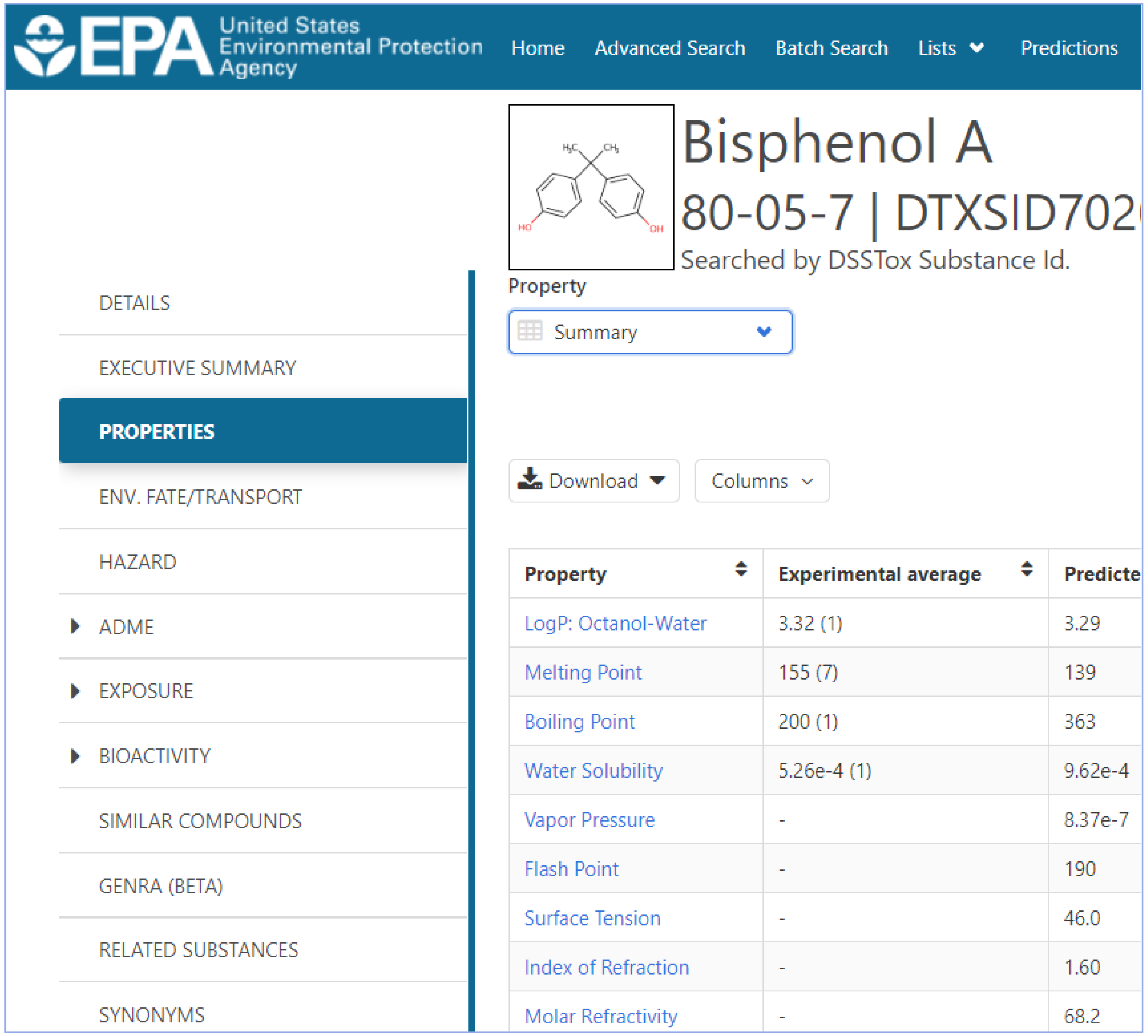
MoleculeID	CATMoS_VT_pred	CATMoS_NT_pred	CATMoS_EPA_pred	CATMoS_GHS_pred	CATMoS_LD50_pred	AD_CATMoS	AD_index_CATMoS	Conf_index_CATMoS
'123-91-1'	0	1	3	5	3.4053	1	1	0.9500
'67-97-0'	1	0	1	2	1.2845	1	1	0.8684

Soon on NTP/ICE and EPA CompTox dashboard

https://ntp.niehs.nih.gov/

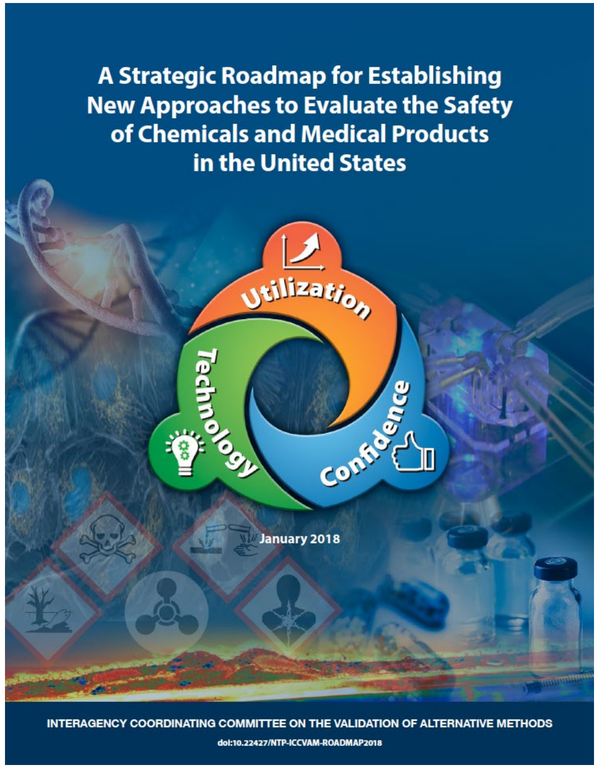
https://comptox.epa.gov/dashboard





The “3C” Concept at Work!

- Success of the project was due in great part to the use of the 3C concept as well as up-front and continuous engagement of regulators in the process



Communication

Collaboration

Commitment

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

Acknowledgements

THANK YOU!

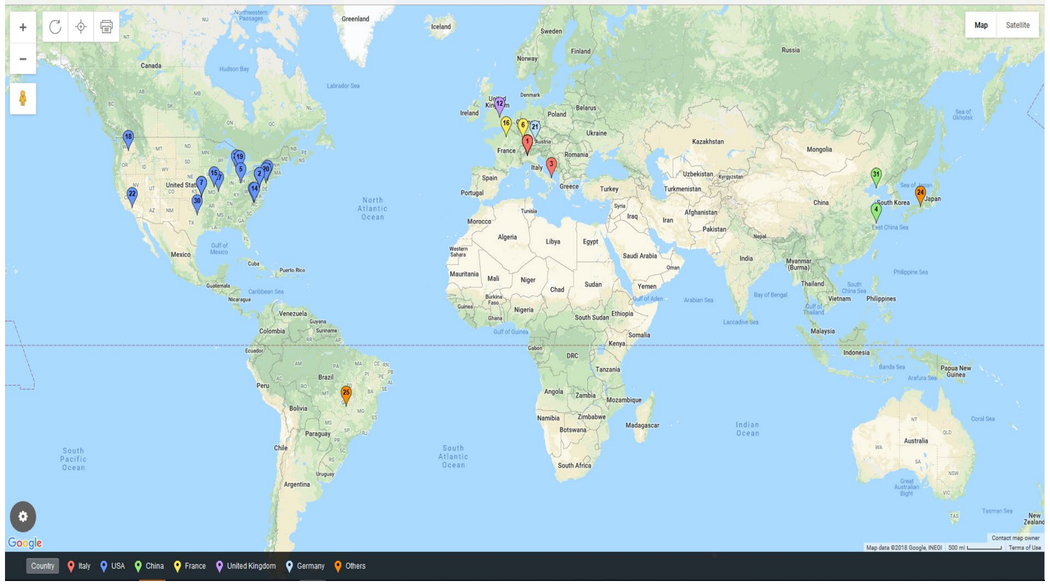
- ICCVAM Acute Toxicity Workgroup

- EPA/NCCT
  - Grace Patlewicz
  - Jeremy Fitzpatrick

- ILS/NICEATM
  - Agnes Karmaus
  - Dave Allen
  - Shannon Bell
  - Patricia Ceger
  - Judy Strickland
  - Amber Daniel

- NTP/NICEATM
  - Nicole Kleinstreuer
  - Warren Casey

All CATMoS international collaborators



Feedback welcome: Kamel Mansouri (kmansouri@ils-inc.com)

Technical support was provided by ILS under NIEHS contract HHSN273201500010C.