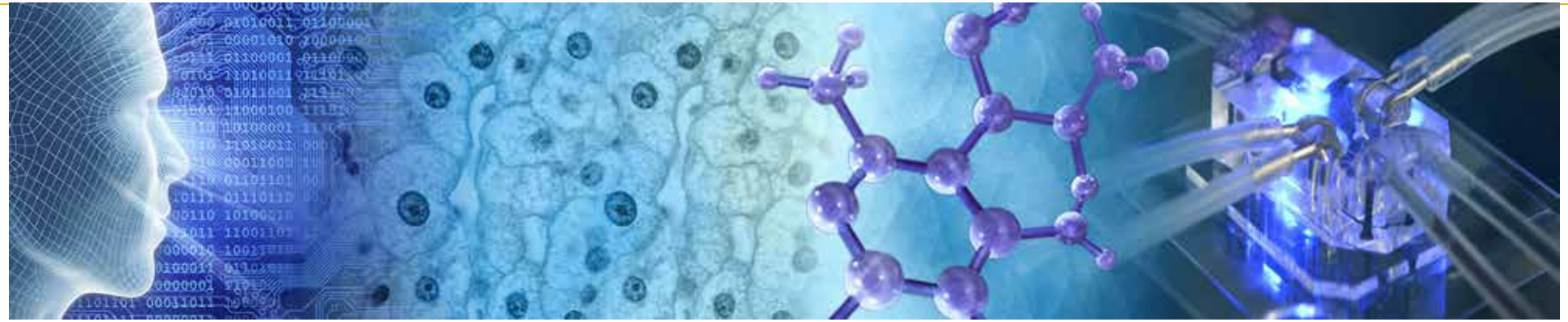
NTP
National Toxicology Program



Collaborative Acute Toxicity Modeling Suite (CATMoS)

Kamel Mansouri
LEAD COMPUTATIONAL CHEMIST, ILS IN SUPPORT OF NICEATM

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)

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)

Hazard

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

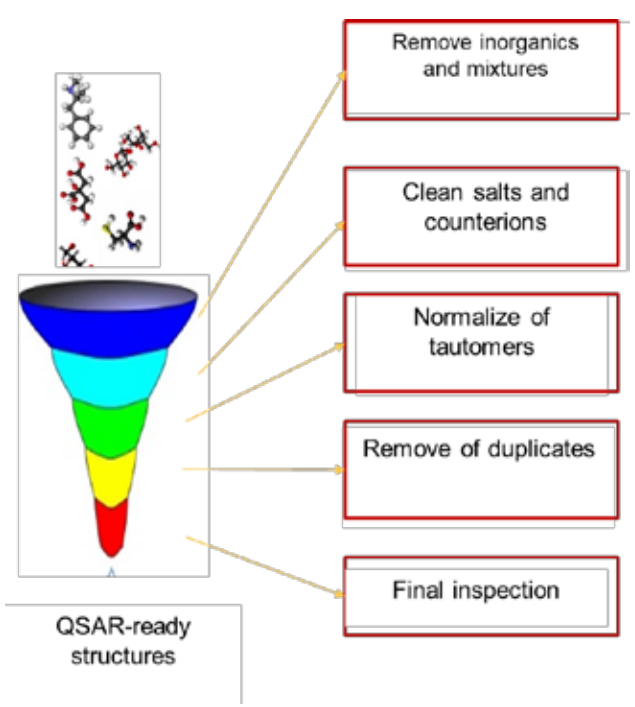
11992 chemicals with accurate structures

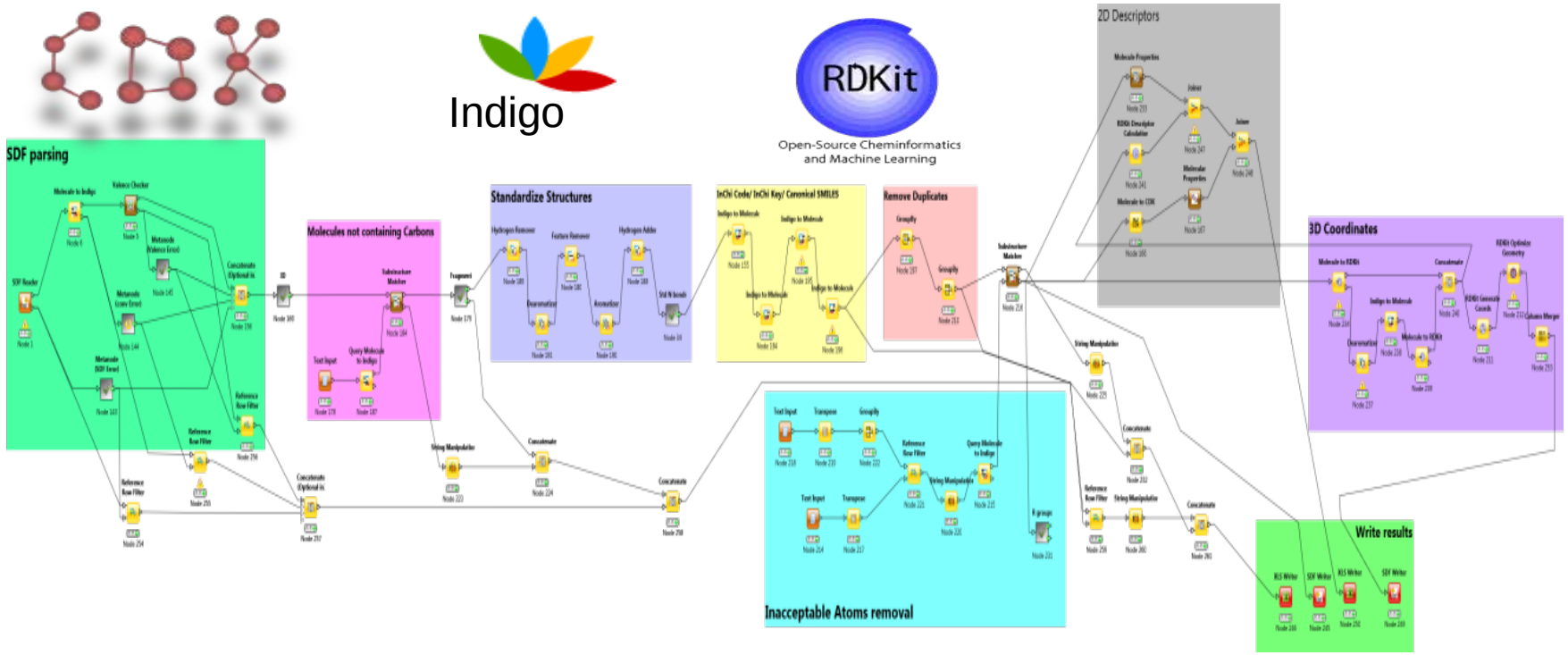
- Very toxic endpoint: 11886 entries (binary, 0/1)
- Non-toxic endpoint: 11871 entries (binary, 0/1)
- EPA endpoint: 11755 entries (categorical, 4 categories)
- GHS endpoint: 11845 entries (categorical, 5 categories)
- LD50 endpoint: 8908 entries (continuous values)

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

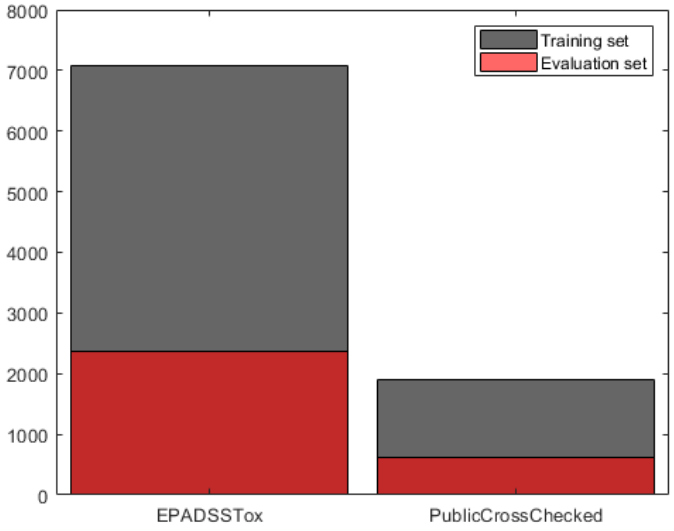
Mansouri et al. (http://ehp.niehs.nih.gov/15-10267/)

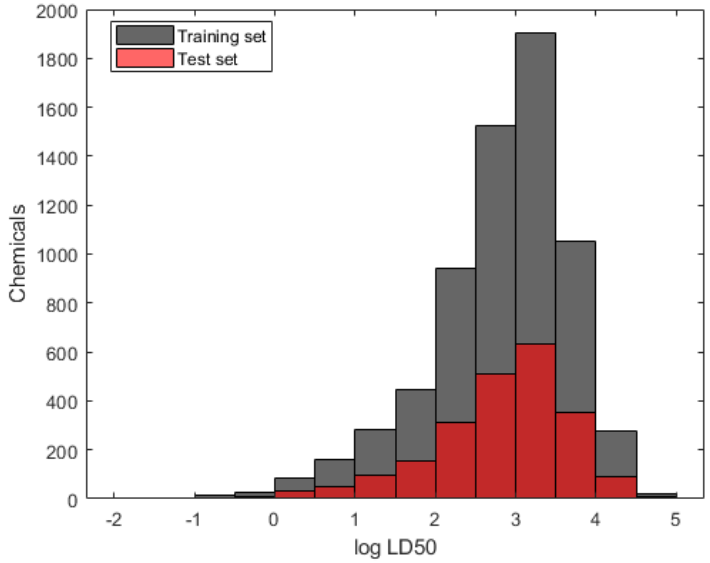
Establishing Modeling Dataset

- **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)

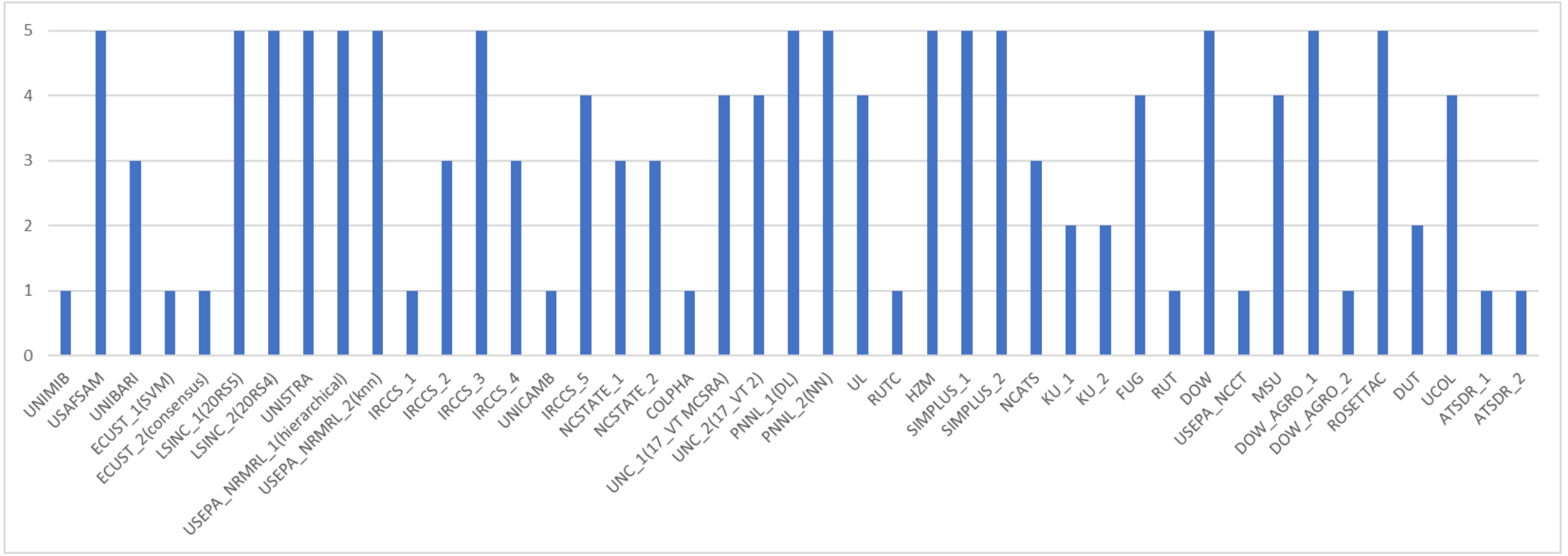
ChemMaps landscape of CATMoS chemicals

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

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 (Tr) statistics
- Predictivity: statistics on the evaluation set
- Robustness: balance between (Goodness of fit) & (Predictivity)

$$S = 0.3 * (Goodness\ of\ fit) + 0.45 * (Predictivity) + 0.25 * (Robustness)$$

Categorical models (binary and multi-class):

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

$$Predictivity = 0.7 * (BA_{Eval}) + 0.3 * \left(1 - |\widetilde{Sn_{Eval} - Sp_{Eval}}|\right)$$

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

$$BA = \frac{(Sn + Sp)}{2} \qquad Sn = \frac{TP}{TP + FN} \qquad Sp = \frac{TN}{TN + FP}$$

Continuous models:

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

$$Predictivity = R^2_{Eval}$$

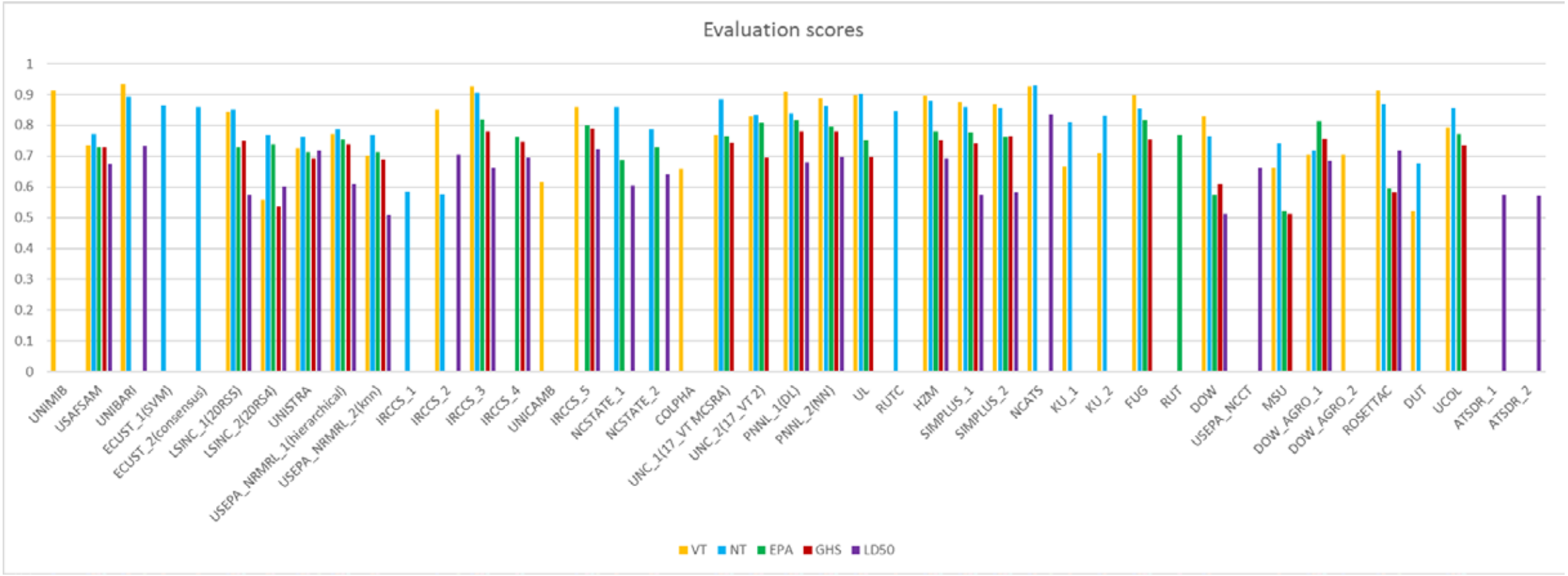
$$Robustness = 1 - |R^2_{Tr} - R^2_{Eval}|$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n_{TR}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_{TR}} (y_i - \bar{y})^2}$$

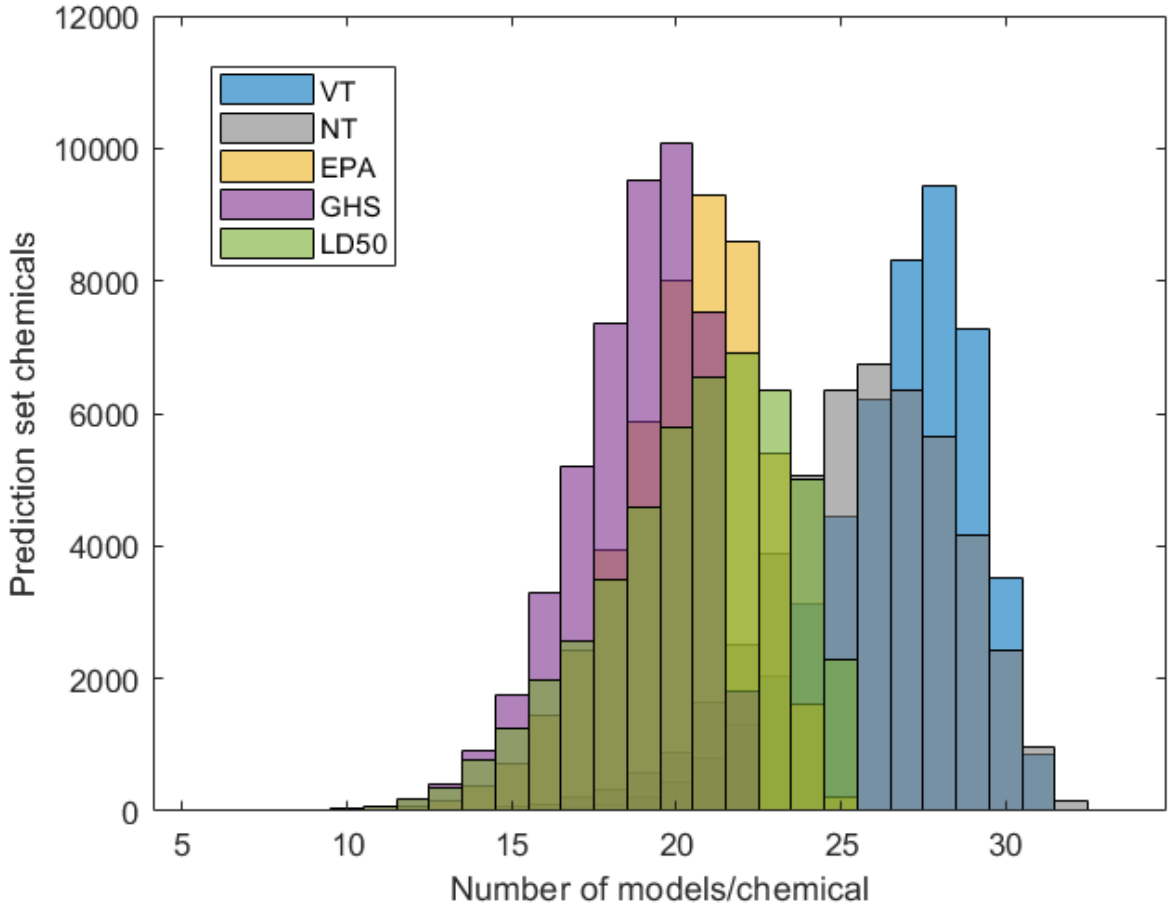
$\hat{y}_i$ and y_i are the estimated and observed responses

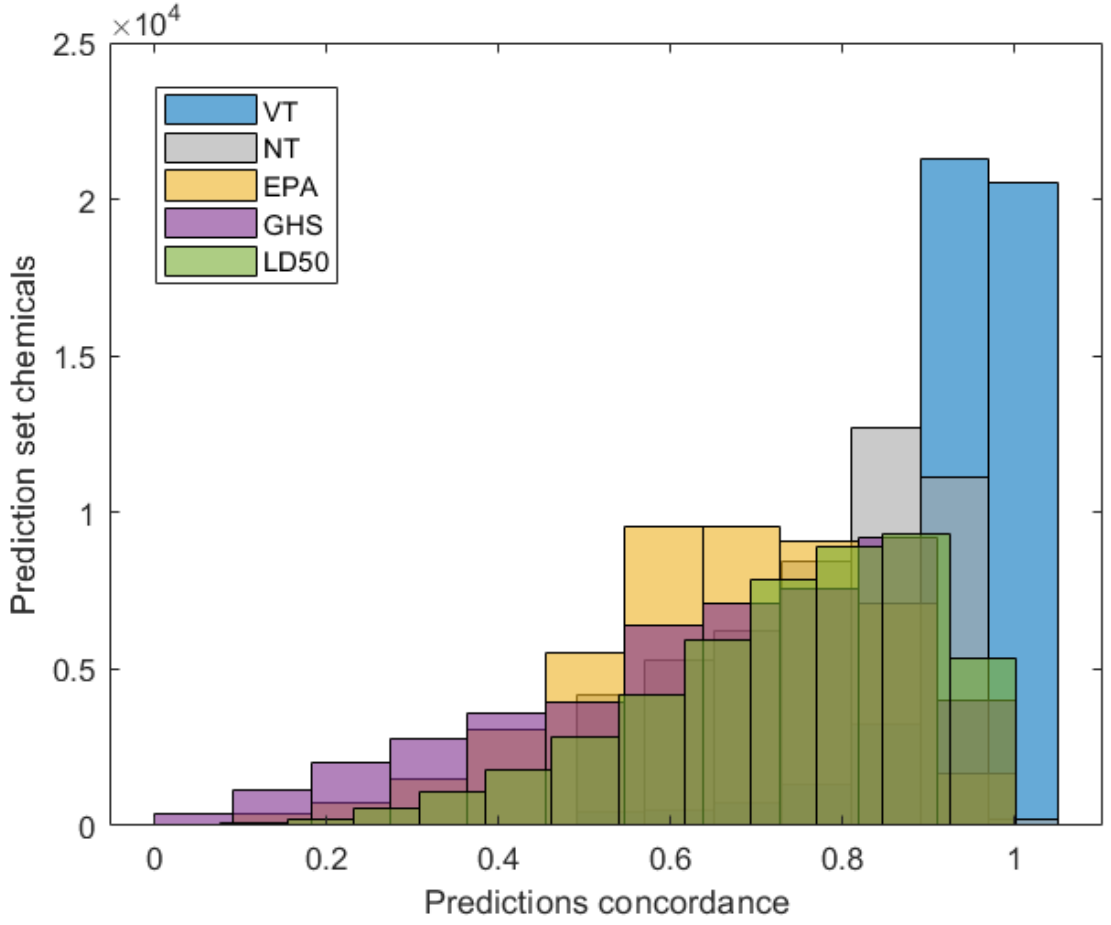
Evaluation results

Quantitative evaluation



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

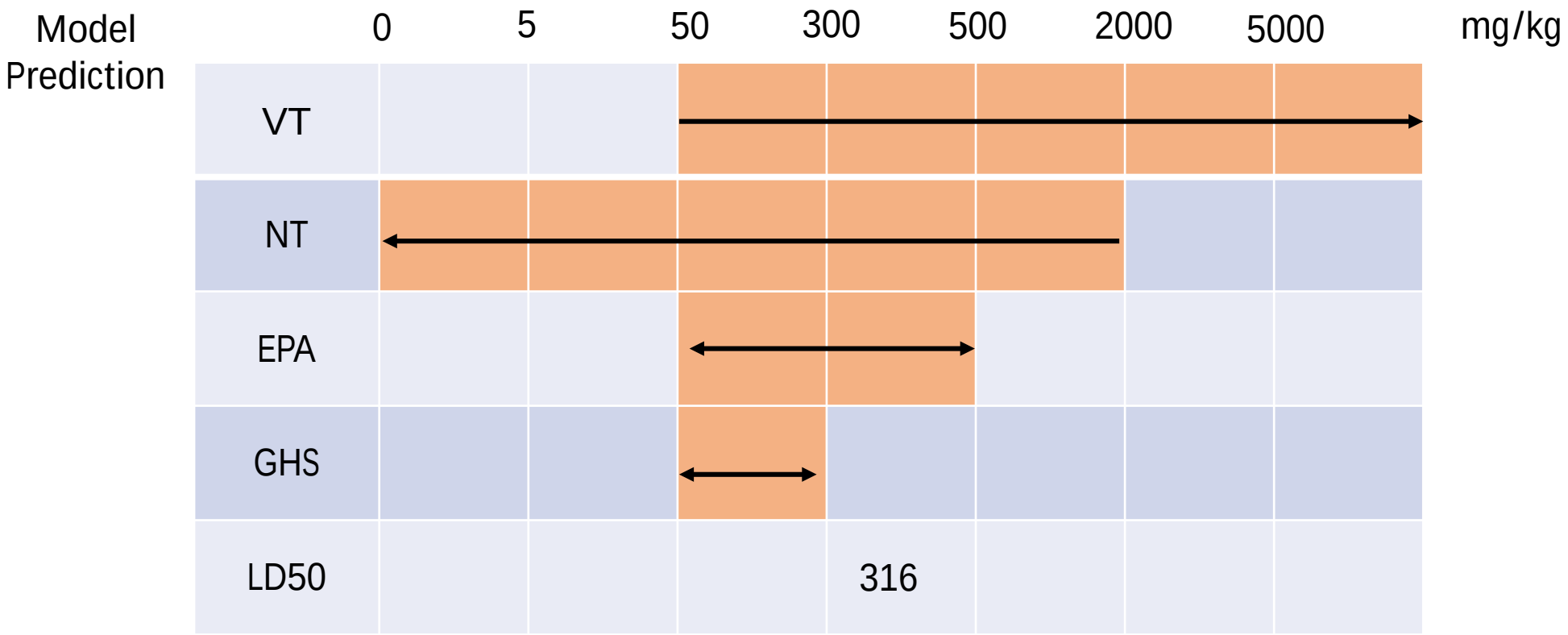
Consistent consensus models/predictions

- VT
- NT
- GHS
- EPA
- LD50

Consensus representing all ~140 models

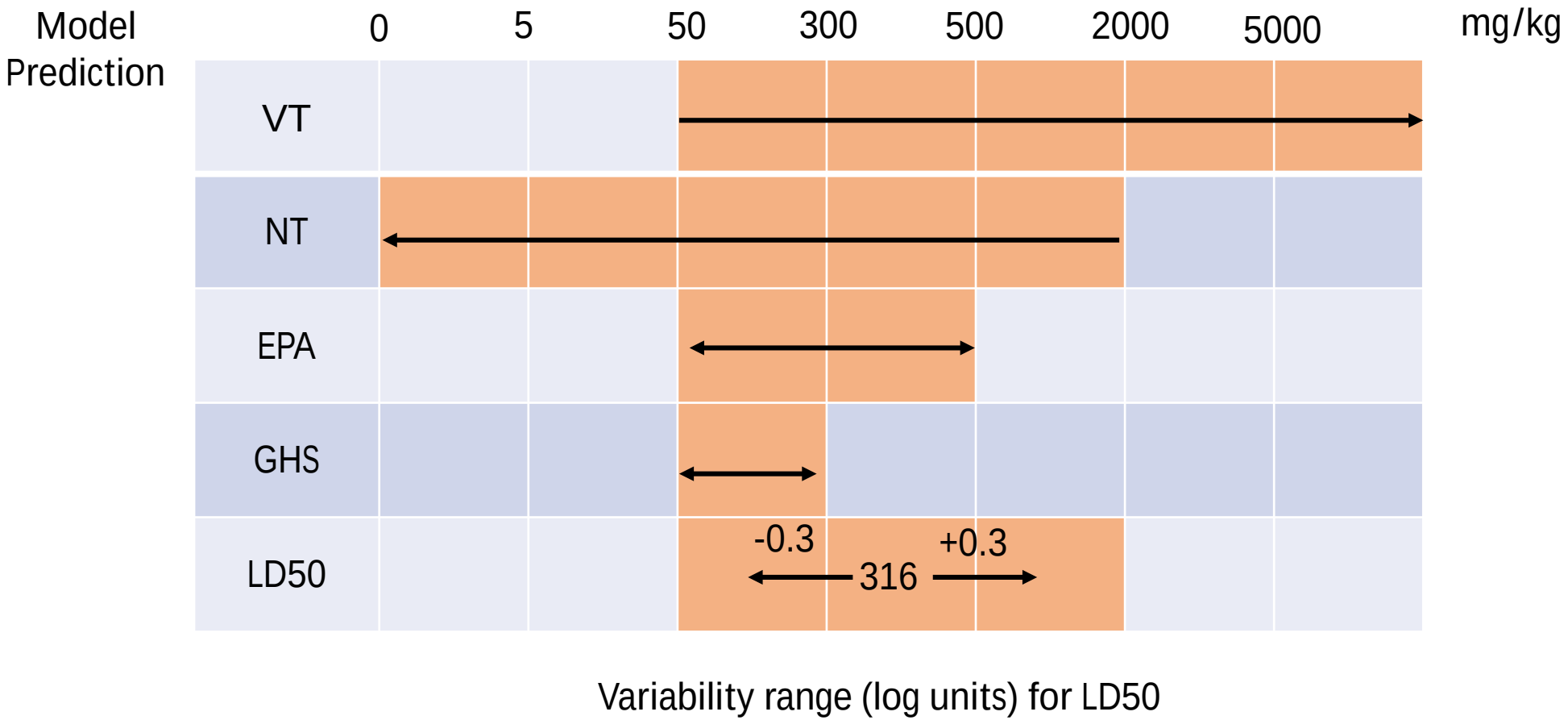
WoE approach to combine the 5 endpoints

	VT	NT	EPA	GHS	LD50
molX	0	0	2	3	2.5



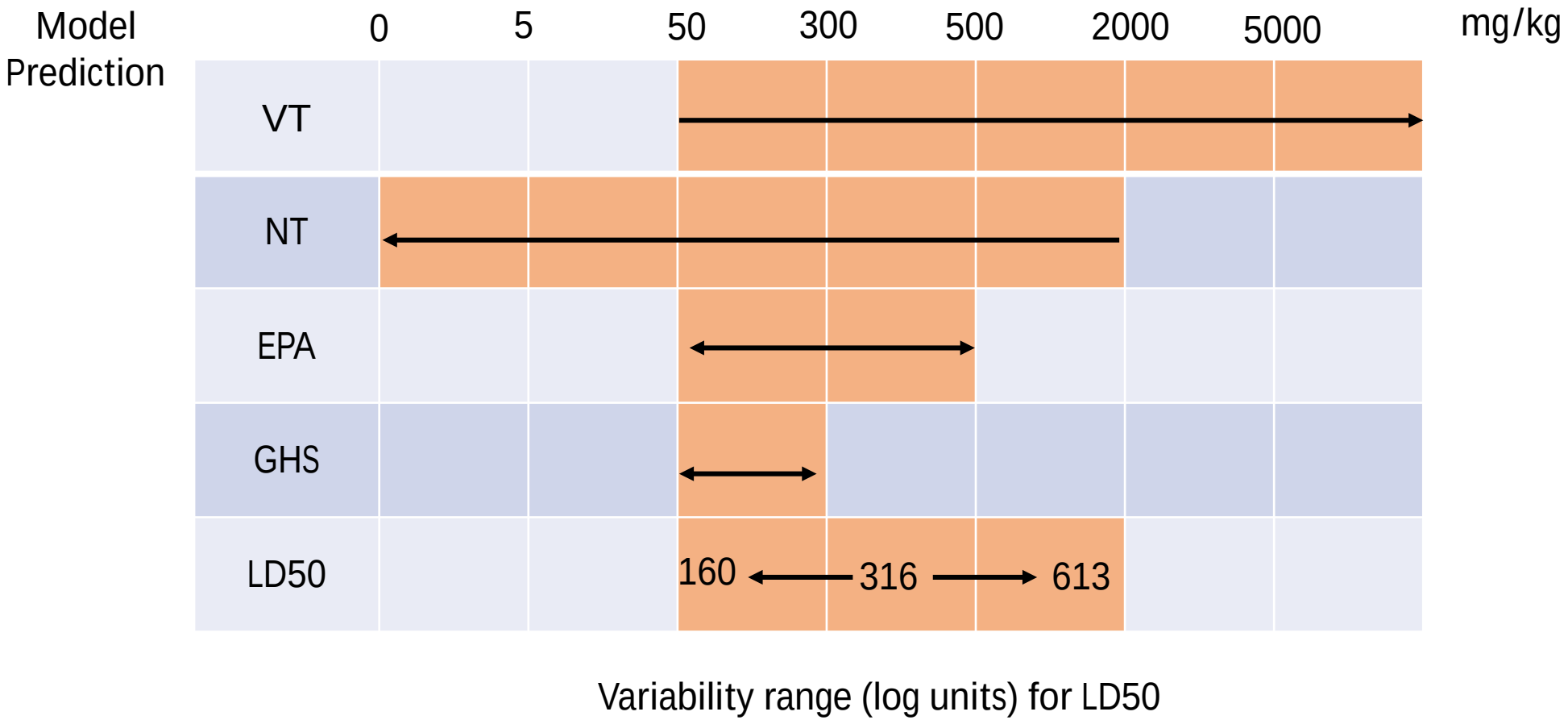
WoE approach to combine the 5 endpoints

	VT	NT	EPA	GHS	LD50
molX	0	0	2	3	2.5



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

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	160 ← 1	1	1 → 613			
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
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 ↔ 613)	1	1			
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

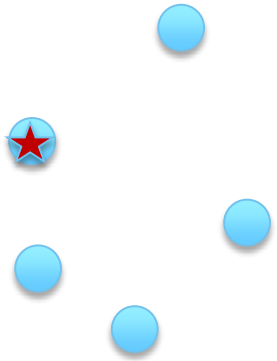
	Very Toxic		Non-Toxic		EPA		GHS	
	Train	Eval	Train	Eval	Train	Eval	Train	Eval
Sensitivity	0.87	0.70	0.88	0.67	0.81	0.62	0.80	0.58
Specificity	0.99	0.97	0.97	0.90	0.92	0.86	0.95	0.90
Balanced Accuracy	0.93	0.84	0.92	0.78	0.87	0.74	0.88	0.74
***In vivo* Balanced Accuracy**	0.81		0.89		0.82		0.79	

	LD50 values		LD50 values
	Train	Eval	***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

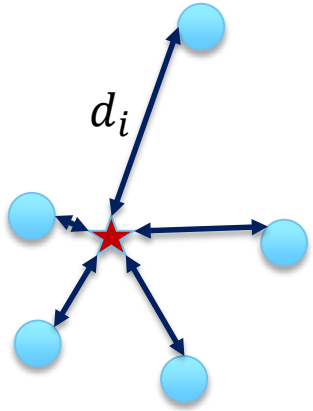
Extended CATMoS predictions

Weighted read-across



$d_1 = 0$

$Pred_i = N_i$



$d_1 \neq 0$

$w_i = f(d_i)$

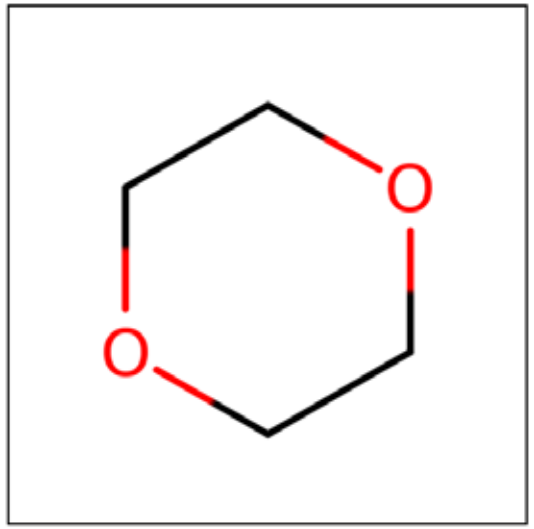
$Pred_i = f(w_i, N_i)$

★ 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

CATMoS prediction examples



United States Environmental Protection Agency

1,4-Dioxane

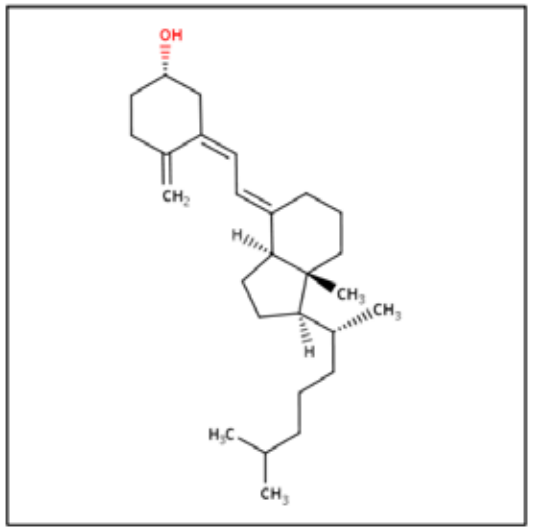
123-91-1 | DTXSID4020533

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



United States Environmental Protection Agency

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

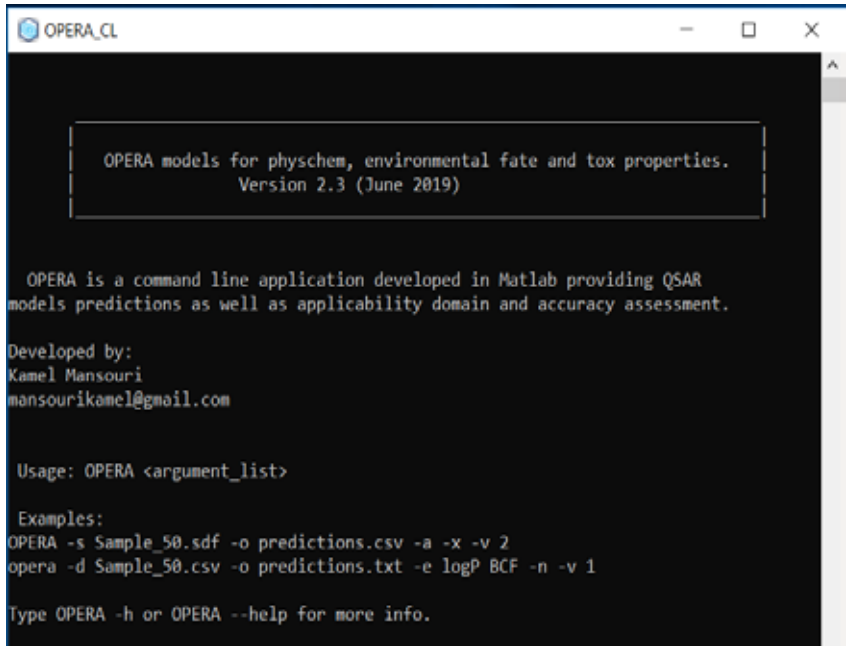
Collaboration with ATWG partners and ICCVAM agencies

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		

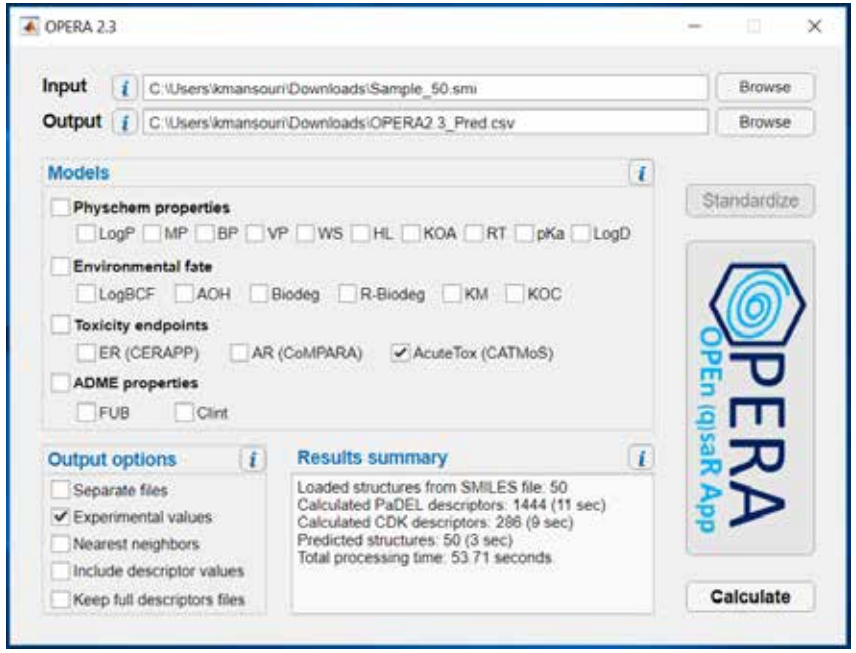
Evaluate and optimize CATMoS predictions based on lists of interest

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

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

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

OPERA2

Since OPERA v1.5

Physchem & Environmental fate:

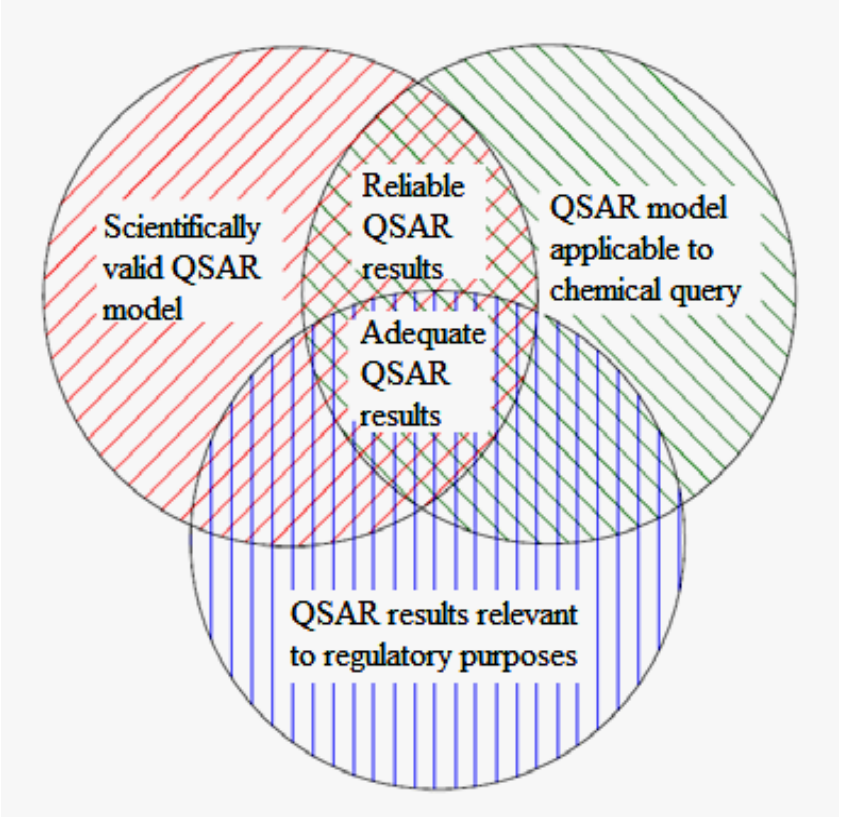
Model	Property
AOH	Atmospheric Hydroxylation Rate
BCF	Bioconcentration Factor
BioHL	Biodegradation Half-life
RB	Ready Biodegradability
BP	Boiling Point
HL	Henry's Law Constant
KM	Fish Biotransformation Half-life
KOA	Octanol/Air Partition Coefficient
LogP	Octanol-water Partition Coefficient
MP	Melting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

New since OPERA 2.0

- 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)**

Inform Regulatory Decisions

Principle	Description
1) A defined endpoint	Any **physicochemical, biological or environmental** effect that can be measured and therefore modelled.
2) An unambiguous algorithm	**Ensure transparency** in the description of the model algorithm.
3) A defined domain of applicability	**Define limitations** in terms of the types of **chemical structures**, physicochemical properties and mechanisms of action for which the models can generate **reliable predictions**.
4) Appropriate measures of goodness-of-fit, robustness and predictivity	a) The internal **fitting** performance of a model b) the **predictivity** of a model, determined by using an appropriate **external test set**.
5) Mechanistic interpretation, if possible	Mechanistic **associations** between the **descriptors** used in a model and the **endpoint being predicted**.

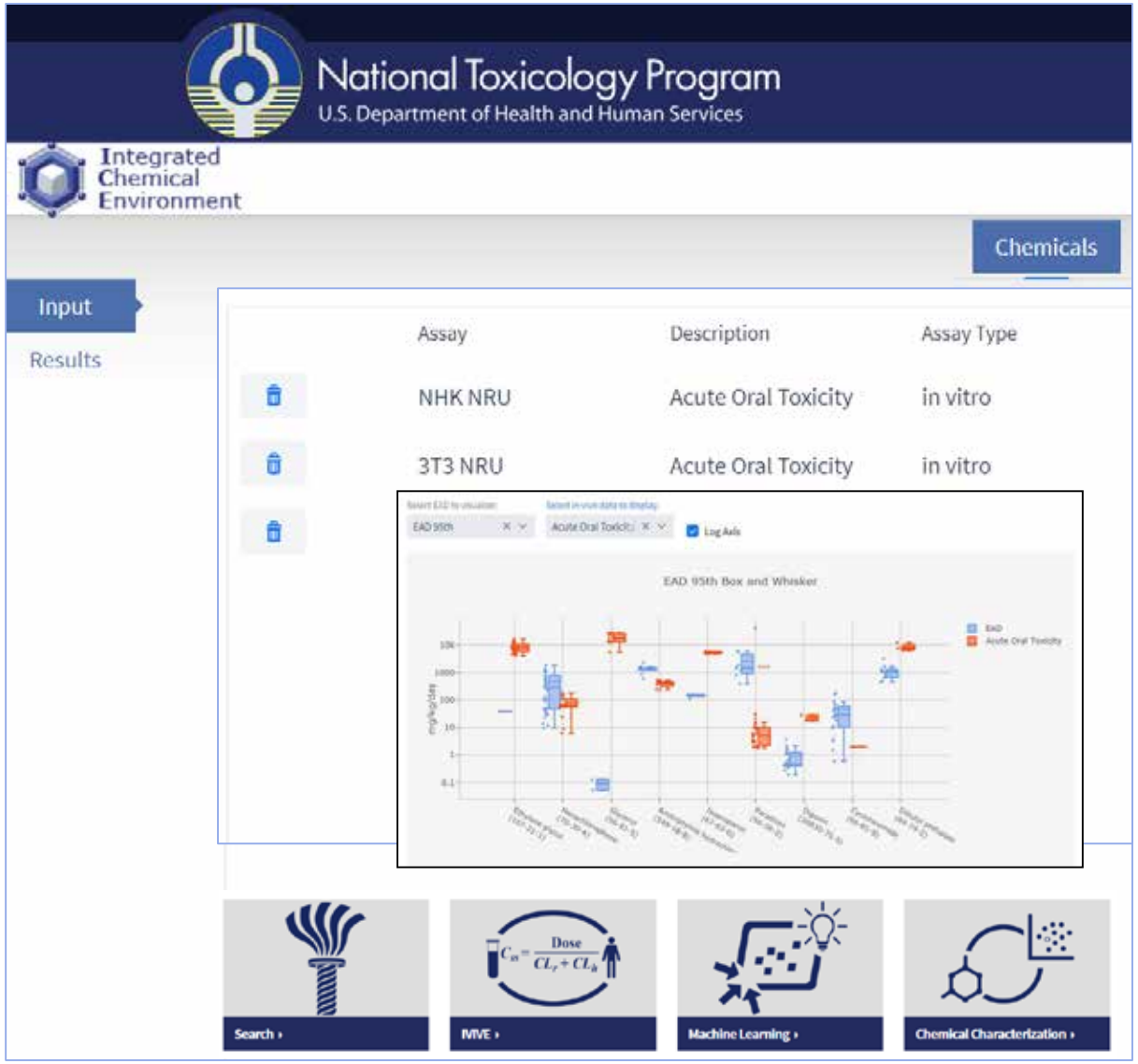


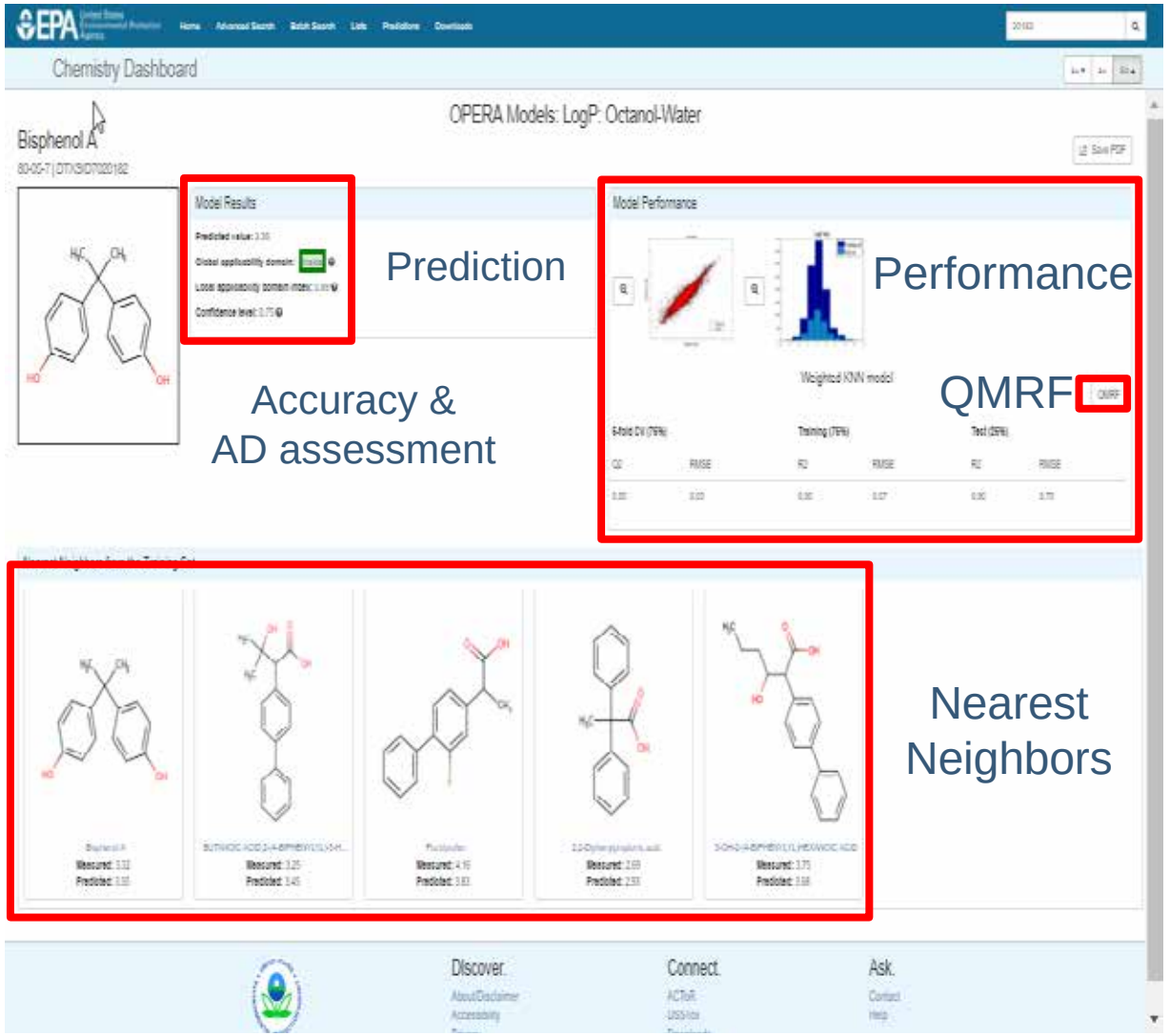
Mansouri et al. J Cheminform (2018). https://doi.org/10.1186/s13321-018-0263-1

Soon on NTP/ICE and EPA CompTox dashboard

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

https://comptox.epa.gov/dashboard



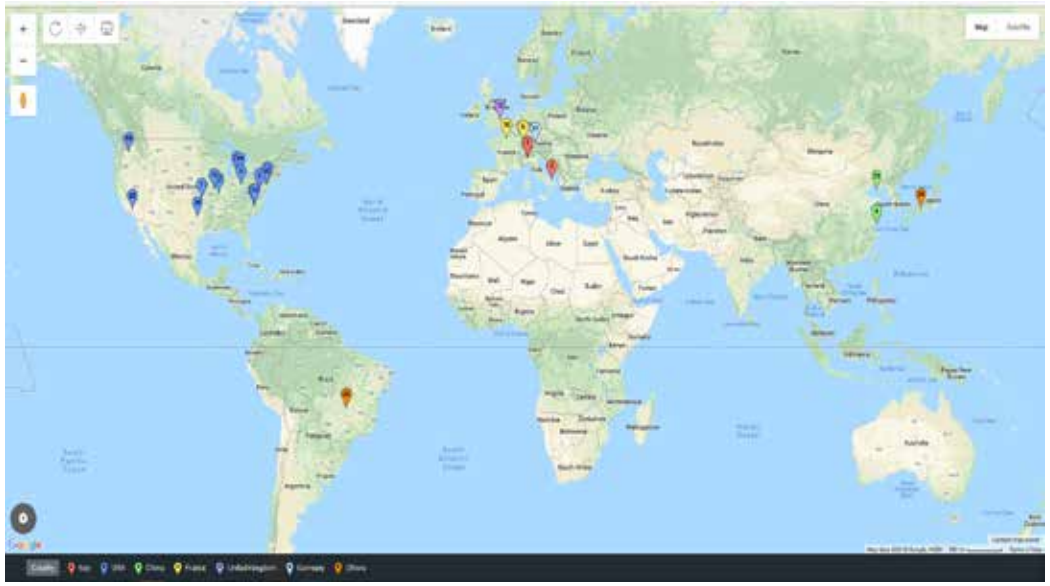


Acknowledgements

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.