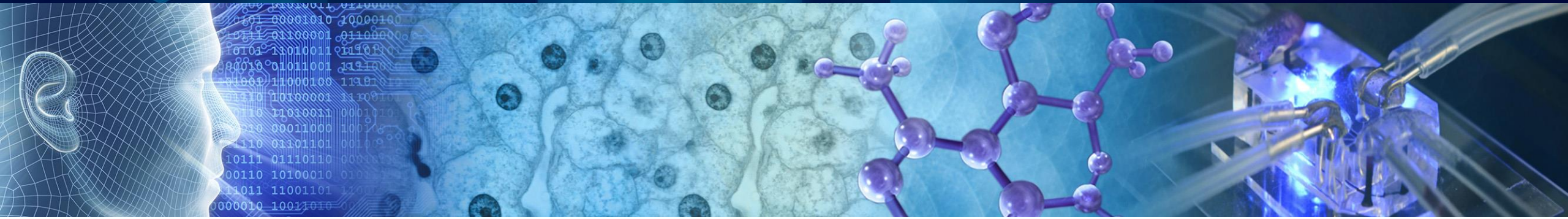
National Institute of
Environmental Health Sciences
Division of Translational Toxicology

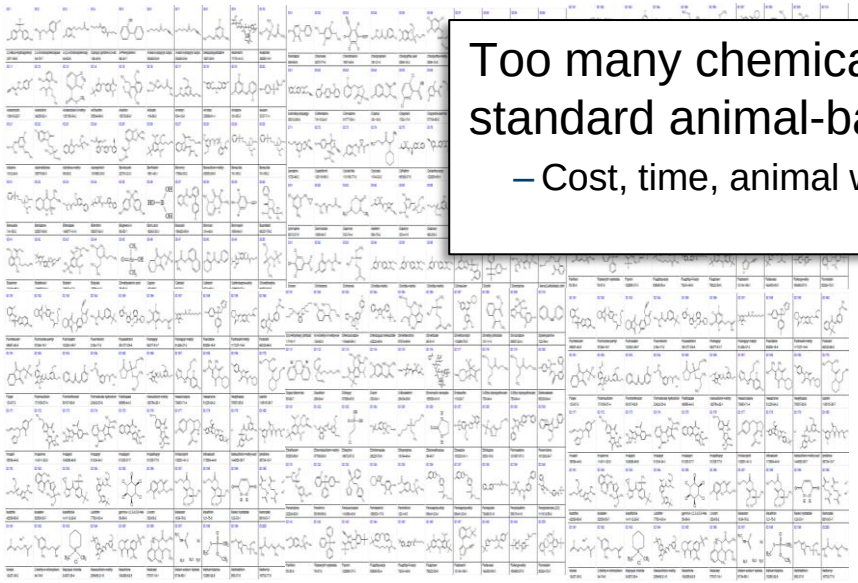


Collaborative modeling project for predicting acute oral toxicity (CATMoS)

Kamel Mansouri
Computational Chemist
NIH/NIEHS/DTT/NICEATM, RTP, NC, USA

The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of any federal agency.

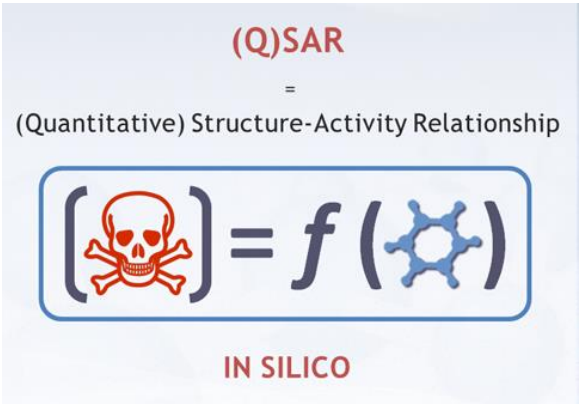
In silico screening



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 wildlife 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**.**

International collaborative projects

CERAPP

Collaborative Estrogen Receptor Activity Prediction Project **(2015/16)**

Mansouri et al. (https://doi.org/10.1289/ehp.1510267)

CoMPARA

Collaborative **M**odeling **P**roject for **A**ndrogen **R**eceptor **A**ctivity **(2017/18)**

Mansouri et al. (https://doi.org/10.1289/EHP5580)

CATMoS

Collaborative **A**cute **T**oxicity **Mo**deling **S**uite **(2018/20)**

Kleinstreuer et al. (https://doi.org/10.1016/j.comtox.2018.08.002)
Mansouri et al. (https://doi.org/10.1289/EHP8495)

Endocrine Disruptor Screening Program

Acute Toxicity Workgroup: alternative methods

ICCVAM: Interagency Coordinating Committee on the Validation of Alternative Methods

National Institute of Environmental Health Sciences
Division of Translational Toxicology

International Consortium

Over 100 scientists from around the globe representing academia, industry, and government contributed

ehp Environmental Health Perspectives

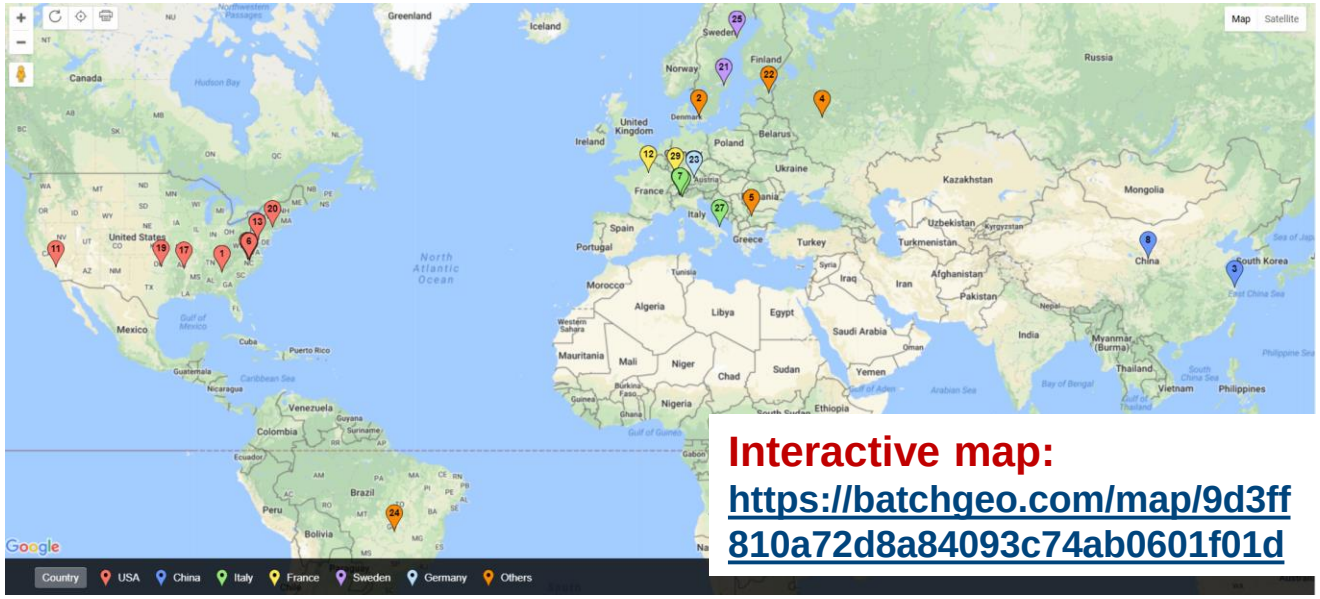
HOME ISSUE IN PROGRESS ARCHIVES COLLECTIONS AUTHORS REVIEWERS ABOUT INTRODUCING *JHP*

Vol. 129, No. 4 | Research

CATMoS: Collaborative Acute Toxicity Modeling Suite

is corrected by

Kamel Mansouri, Agnes L. Karmaus, Jeremy Fitzpatrick, Grace Patlewicz, Prachi Pradeep, Domenico Alberga, Nathalie Alepee, Timothy E.H. Allen, Dave Allen, Vinicius M. Alves, Carolina H. Andrade, Tyler R. Auernhammer, Davide Ballabio, Shannon Bell, Emilio Benfenati, Sudin Bhattacharya, Joyce V. Bastos, Stephen Boyd, J.B. Brown, Stephen J. Capuzzi, Yaroslav Chushak, Heather Ciallella, Alex M. Clark, Viviana Consonni, Pankaj R. Daga, Sean Ekins, Sherif Farag, Maxim Fedorov, Denis Fourches, Domenico Gadaleta, Feng Gao, Jeffery M. Gearhart, Garett Goh, Jonathan M. Goodman, Francesca Grisoni, Christopher M. Grulke, Thomas Hartung, Matthew Hirn, Pavel Karpov, Alexandru Korotcov, Giovanna J. Lavado, Michael Lawless, Xinhao Li, Thomas Luechtefeld, Filippo Lunghini, Giuseppe F. Mangiatordi, Gilles Marcou, Dan Marsh, Todd Martin, Andrea Mauri, Eugene N. Muratov, Glenn J. Myatt, Dac-Trung Nguyen, Orazio Nicolotti, Reine Note, Paritosh Pande, Amanda K. Parks, Tyler Peryea, Ahsan H. Polash, Robert Rallo, Alessandra Roncaglioni, Craig Rowlands, Patricia Ruiz, Daniel P. Russo, Ahmed Sayed, Risa Sayre, Timothy Sheils, Charles Siegel, Arthur C. Silva, Anton Simeonov, Sergey Sosnin, Noel Southall, Judy Strickland, Yun Tang, Brian Teppen, Igor V. Tetko, Dennis Thomas, Valery Tkachenko, Roberto Todeschini, Cosimo Toma, Ignacio Tripodi, Daniela Trisciuzzi, Alexander Tropsha, Alexandre Varnek, Kristijan Vukovic, Zhongyu Wang, Liguo Wang, Katrina M. Waters, Andrew J. Wedlake, Sanjeeva J. Wijeyesakere, Dan Wilson, Zijun Xiao, Hongbin Yang, Gergely Zahoranszky-Kohalmi, Alexey V. Zakharov, Fagen F. Zhang, Zhen Zhang, Tongan Zhao, Hao Zhu, Kimberley M. Zorn, Warren Casey, and Nicole C. Kleinstreuer



Interactive map:
https://batchgeo.com/map/9d3ff810a72d8a84093c74ab0601f01d

National Institute of Environmental Health Sciences
Division of Translational Toxicology

Acute Oral Toxicity: CATMoS

- ICCVAM is developing alternative test methods for the EPA's six pack tests: Acute oral, dermal, inhalation, eye & skin irritation and skin sensitization
- 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

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]

National Institute of Environmental Health Sciences
Division of Translational Toxicology

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®

National Institute of Environmental Health Sciences
Division of Translational Toxicology

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

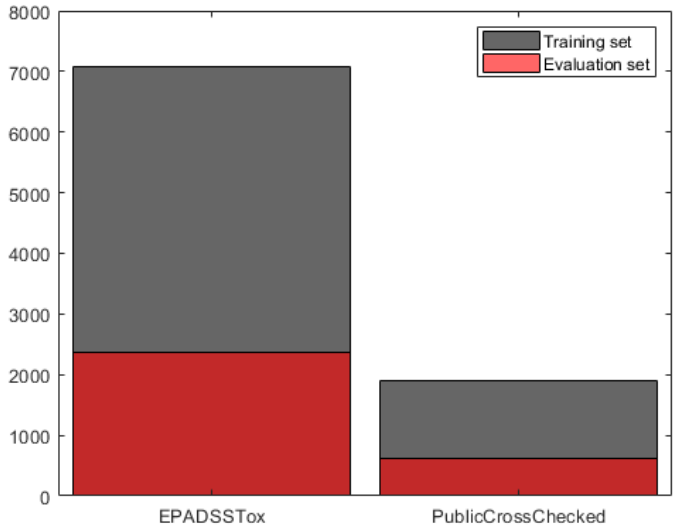
11992 chemicals with accurate structures

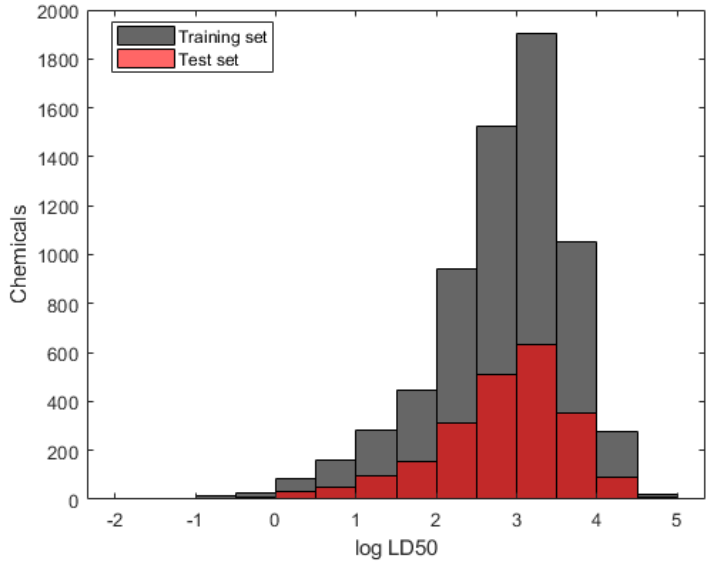
- Very toxic endpoint: 11886 entries (binary, 0/1)
- Non-toxic endpoint: 11871 entries (binary, 0/1)
- EPA endpoint: 11755 entries (categorical, 4 categories)
- GHS endpoint: 11845 entries (categorical, 5 categories)
- LD50 endpoint: 8908 entries (continuous values)

National Institute of Environmental Health Sciences
Division of Translational Toxicology

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





National Institute of Environmental Health Sciences
Division of Translational Toxicology

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)

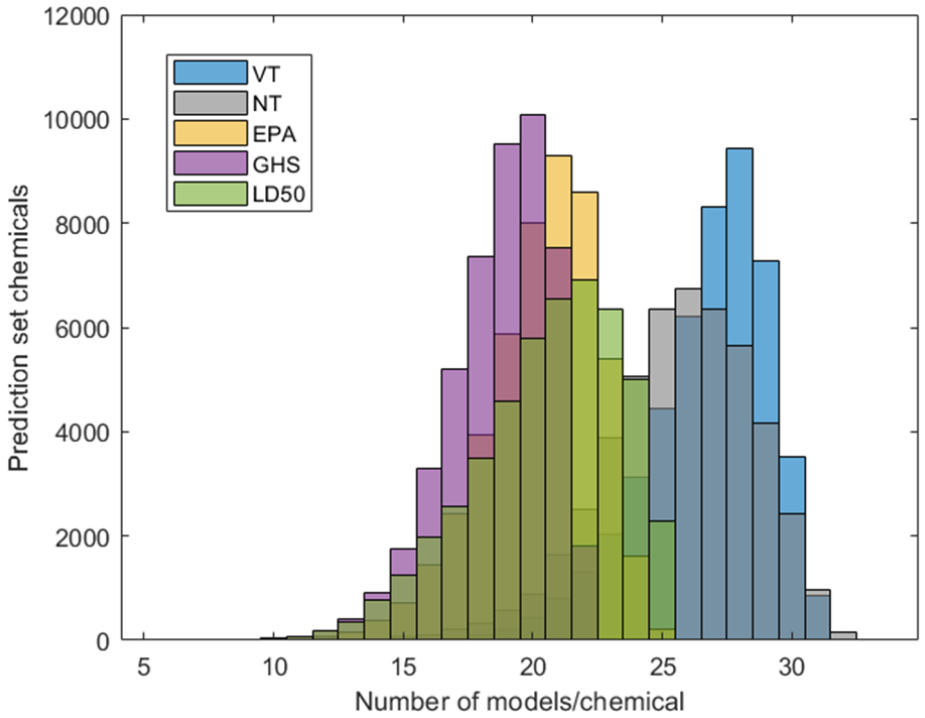
National Institute of Environmental Health Sciences
Division of Translational Toxicology

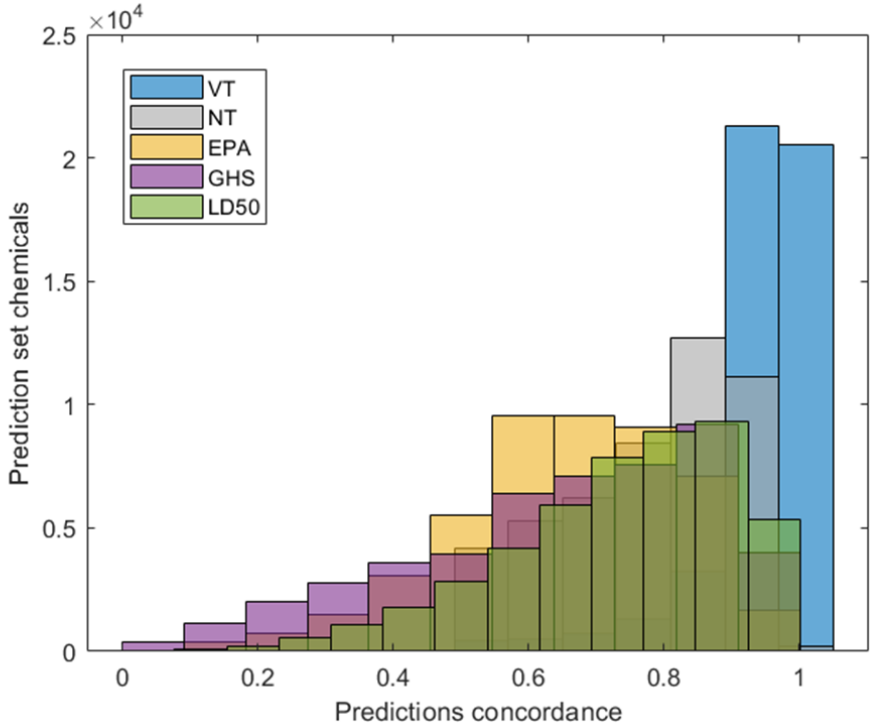
Coverage and concordance of the models

Consortium Comprised 35 Participants/Groups

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

Total: 139 models

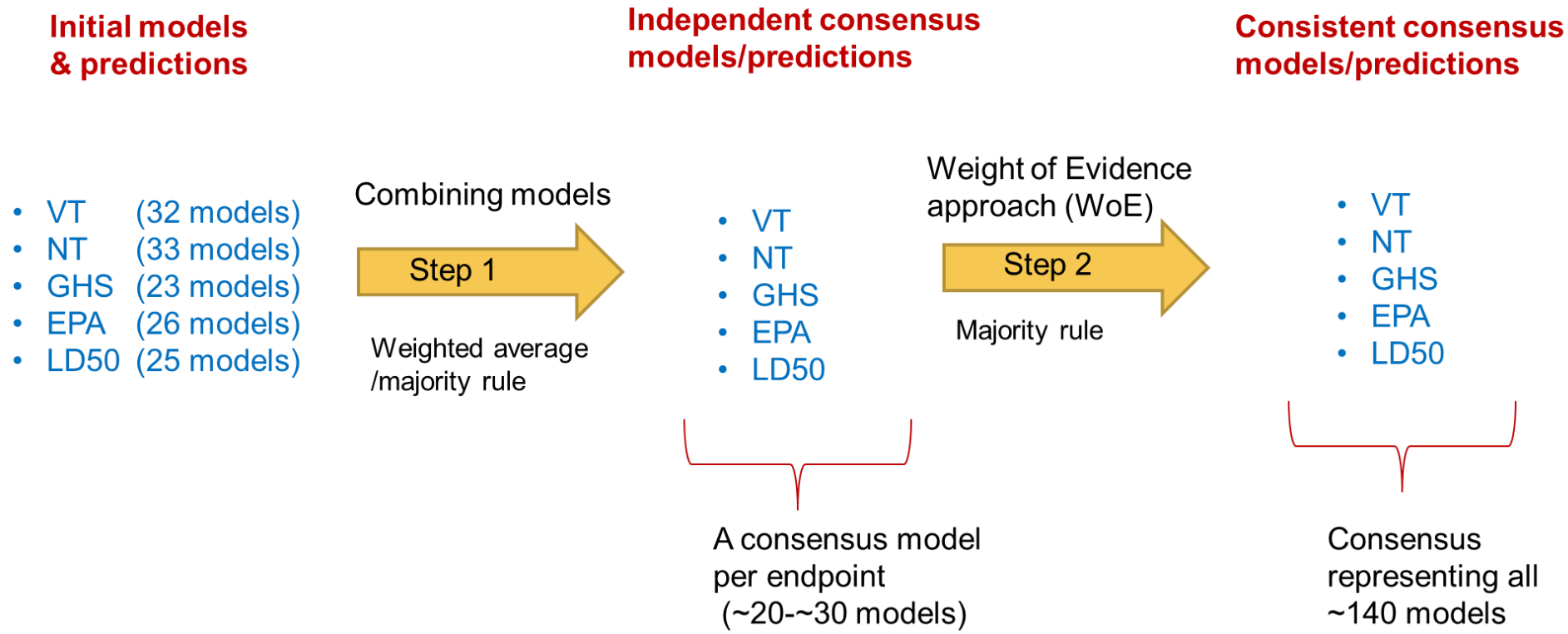




National Institute of Environmental Health Sciences
Division of Translational Toxicology

CATMoS consensus modeling

Steps of combining the single models into consensus



Learn more:

https://www.piscltd.org.uk/wp-content/uploads/2020/01/2020.01.22_CATMoS_Webinar.pdf
https://youtu.be/KjbTnfRTY-0

National Institute of Environmental Health Sciences
Division of Translational Toxicology

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

National Institute of Environmental Health Sciences
Division of Translational Toxicology

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	EPA EFED	160
DOT	3671	FDA CFSAN	22

ELSEVIER

Regulatory Toxicology and Pharmacology
Volume 149, May 2024, 105614

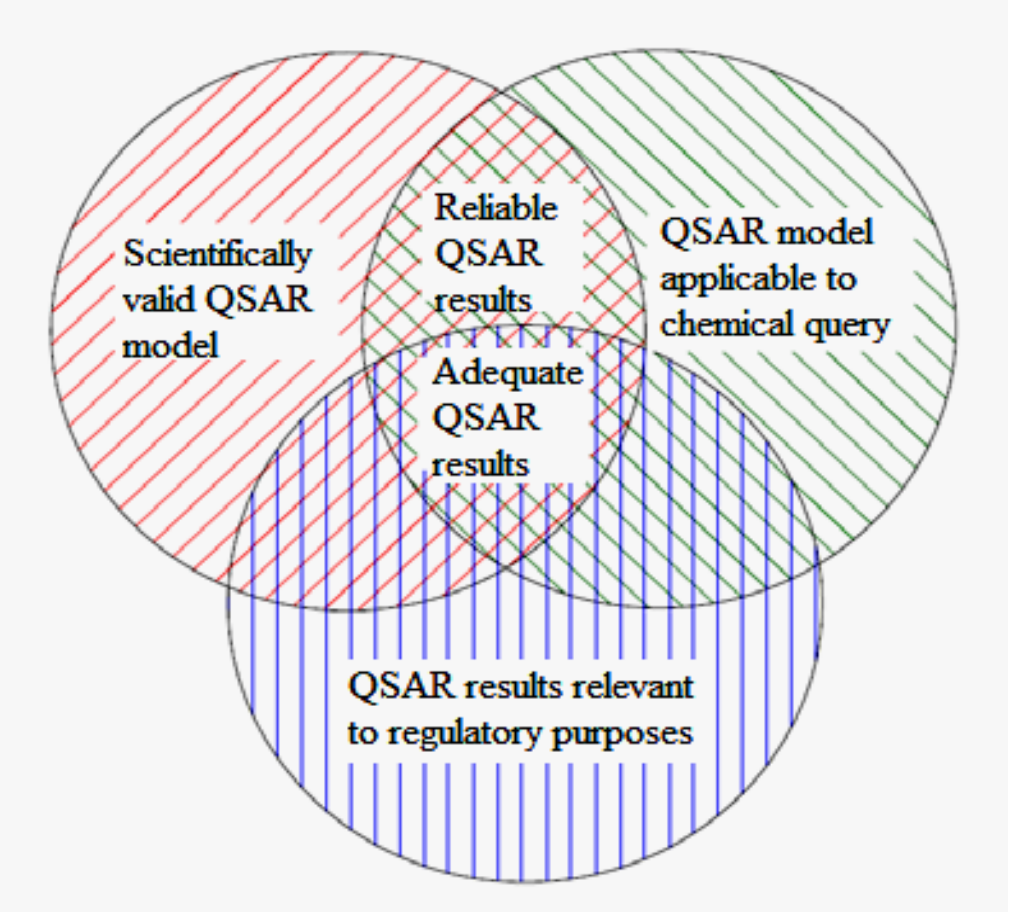
Evaluation of *in silico* model predictions for mammalian acute oral toxicity and regulatory application in pesticide hazard and risk assessment

Patricia L. Bishop [a], Kamel Mansouri [b], William P. Eckel [c], Michael B. Lowit [c], David Allen [d 1], Amy Blankinship [c], Anna B. Lowit [e], D. Ethan Harwood [c], Tamara Johnson [c], Nicole C. Kleinstreuer [b]

OPERA

QSARs for regulatory purposes

The 5 OECD Principles
1) A defined endpoint
2) An unambiguous algorithm
3) A defined domain of applicability
4) Appropriate measures of goodness-of-fit, robustness and predictivity
5) Mechanistic interpretation, if possible



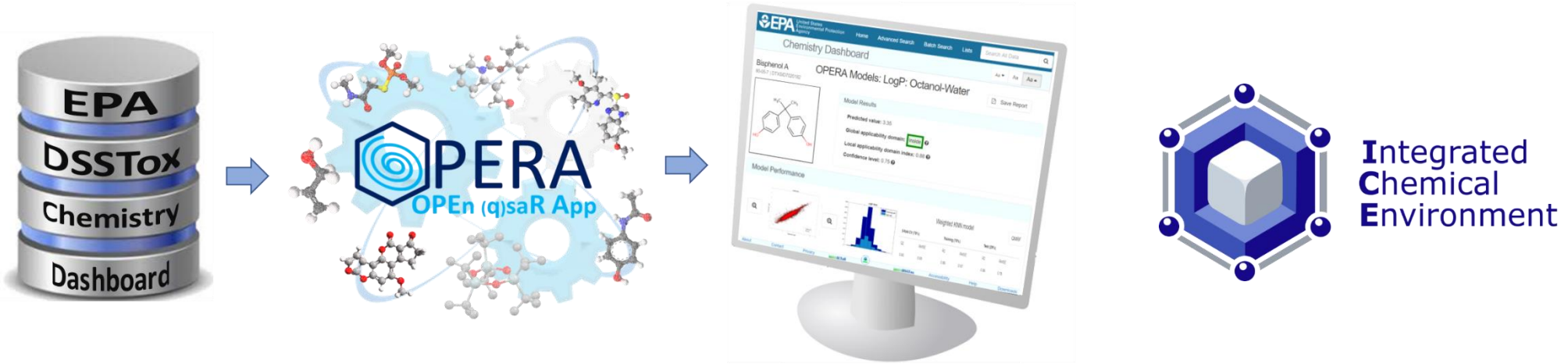
Journal of Cheminformatics
December 2018, 10:10 | Cite as

OPERA models for predicting physicochemical properties and environmental fate endpoints

OPERA
OPEn (q)saR App

OPERA approach

- Curated **open** access datasets (https://doi.org/10.1186/s13321-018-0263-1)
- **Open-source** code (github.com/NIEHS/OPERA)
- **Transparent** unambiguous algorithms (https://qsardb.jrc.ec.europa.eu/qmrf/)
- **Transparent** validated performances (https://doi.org/10.1080/1062936X.2016.1253611)
- **Defined** applicability domain and limitations of the models
- Predictions **available** through:
  - NICEATM's Integrated Chemical Environment (https://ice.ntp.niehs.nih.gov/)
  - The EPA's CompTox Dashboard (https://comptox.epa.gov/dashboard)
  - Free and open-source standalone application (github.com/NIEHS/OPERA)



National Institute of Environmental Health Sciences
Division of Translational Toxicology

OPERA as a standalone desktop application

OPERA standalone application:

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

OPERA models:

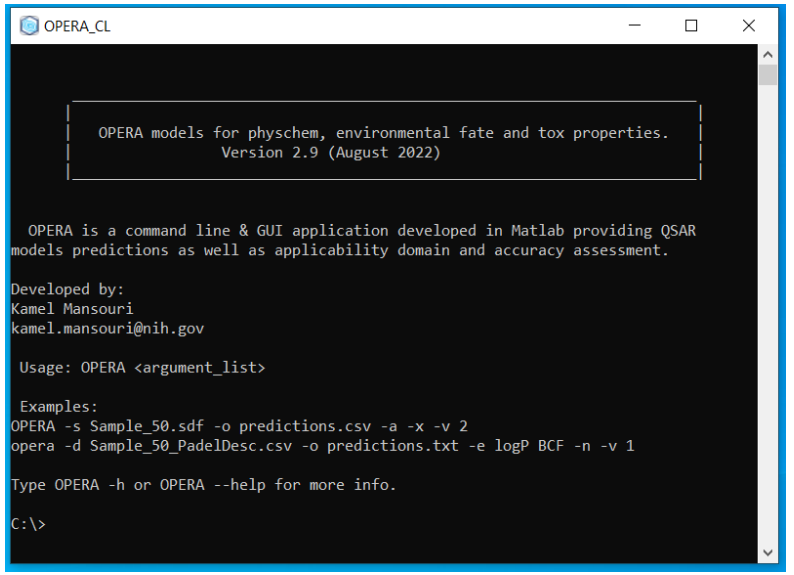
- Physicochemical properties
- Environmental fate
- ADME properties
- Toxicity endpoints

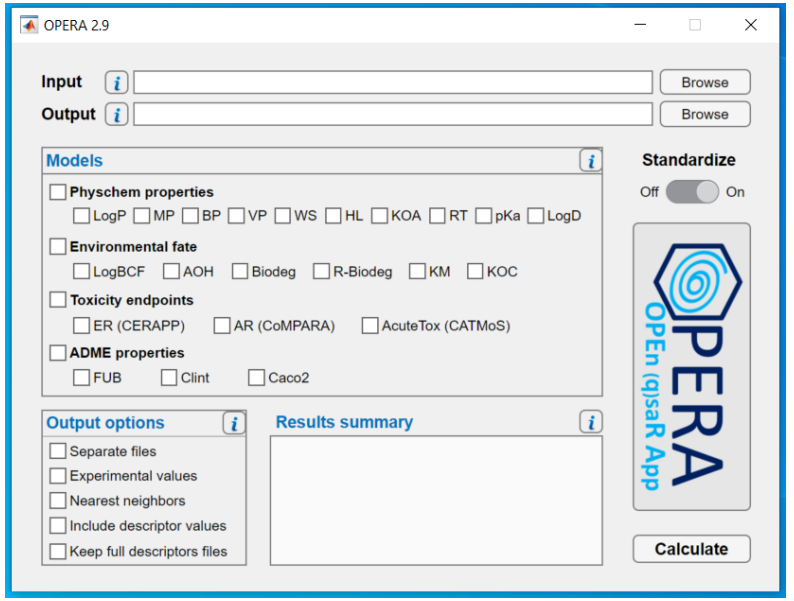
Input options:

- Structure IDs (CAS, DTXSID, InChIKey)
- Structure files (SMILES, SDF, Mol)

Links:

https://github.com/NIEHS/OPERA
https://ntp.niehs.nih.gov/go/opera
https://jcheminf.biomedcentral.com/articles/10.1186/s13321-018-0263-1





National Institute of Environmental Health Sciences
Division of Translational Toxicology

OPERA models (version 2.9)

Physchem properties		Chemicals	Version
BP	**Boiling Point**	**7860**	**2.9**
HL	**Henry's Law Constant**	**2233**	**2.9**
LogP	**Octanol-water Partition Coefficient**	**18154**	**2.9**
MP	**Melting Point**	**22554**	**2.9**
VP	**Vapor Pressure**	**6764**	**2.9**
WS	**Water Solubility**	**9943**	**2.9**
pKa	Acid Dissociation Constant	6503	2.6
KOA	Octanol/Air Partition Coefficient	270	2.6

Environmental fate		Chemicals	Version
AOH	Atmospheric Hydroxylation Rate	692	2.6
BCF	Bioconcentration Factor	626	2.6
BioHL	Biodegradation Half-life	150	2.6
RB	Ready Biodegradability	1603	2.6
KM	Fish Biotransformation Half-life	541	2.6
KOC	Soil Adsorption Coefficient	728	2.6

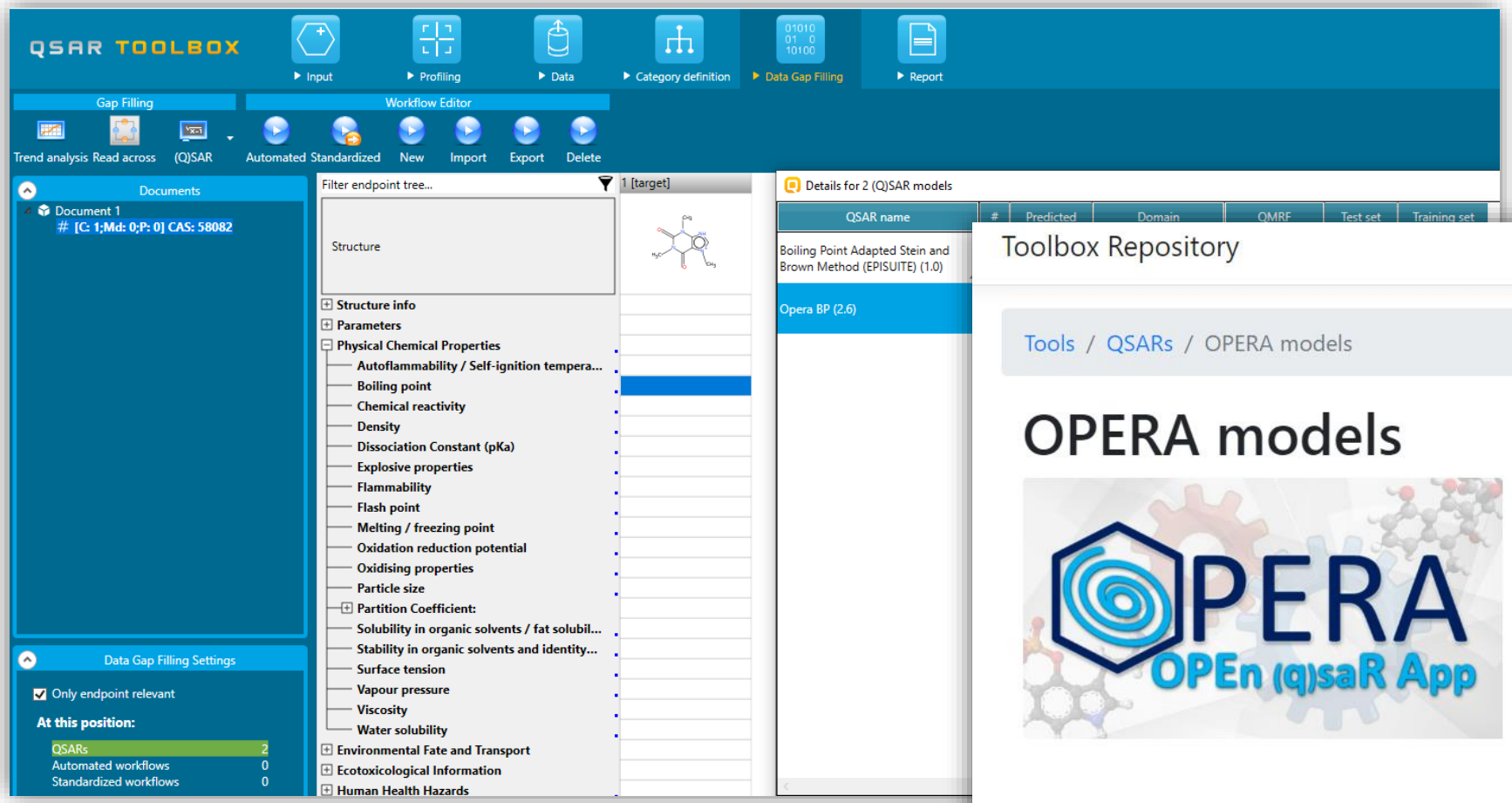
Toxicity endpoints		Chemicals	Version
ER	Estrogen Receptor Activity	32464	2.6
AR	Androgen Receptor Activity	47673	2.6
AcuteTox	Acute Oral Systemic Toxicity	50660	2.6

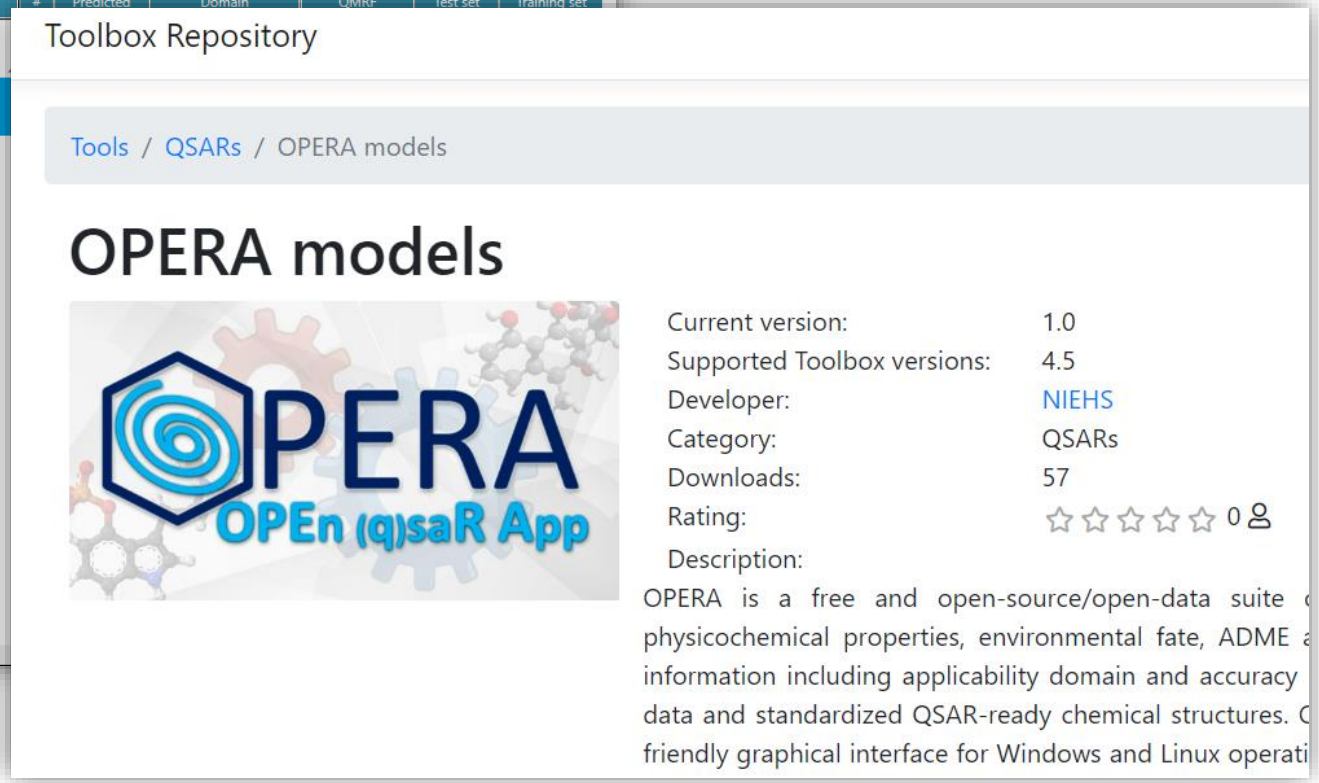
ADME properties		Chemicals	Version
FUB	**Fraction unbound**	**3229**	**2.8**
Clint	**Intrinsic clearance**	**1346**	**2.8**
CACO2	**Caco-2 permeability**	**4601**	**2.8**

New/updated since 2021

National Institute of Environmental Health Sciences
Division of Translational Toxicology

OPERA in the OECD Toolbox





https://repository.qsartoolbox.org/

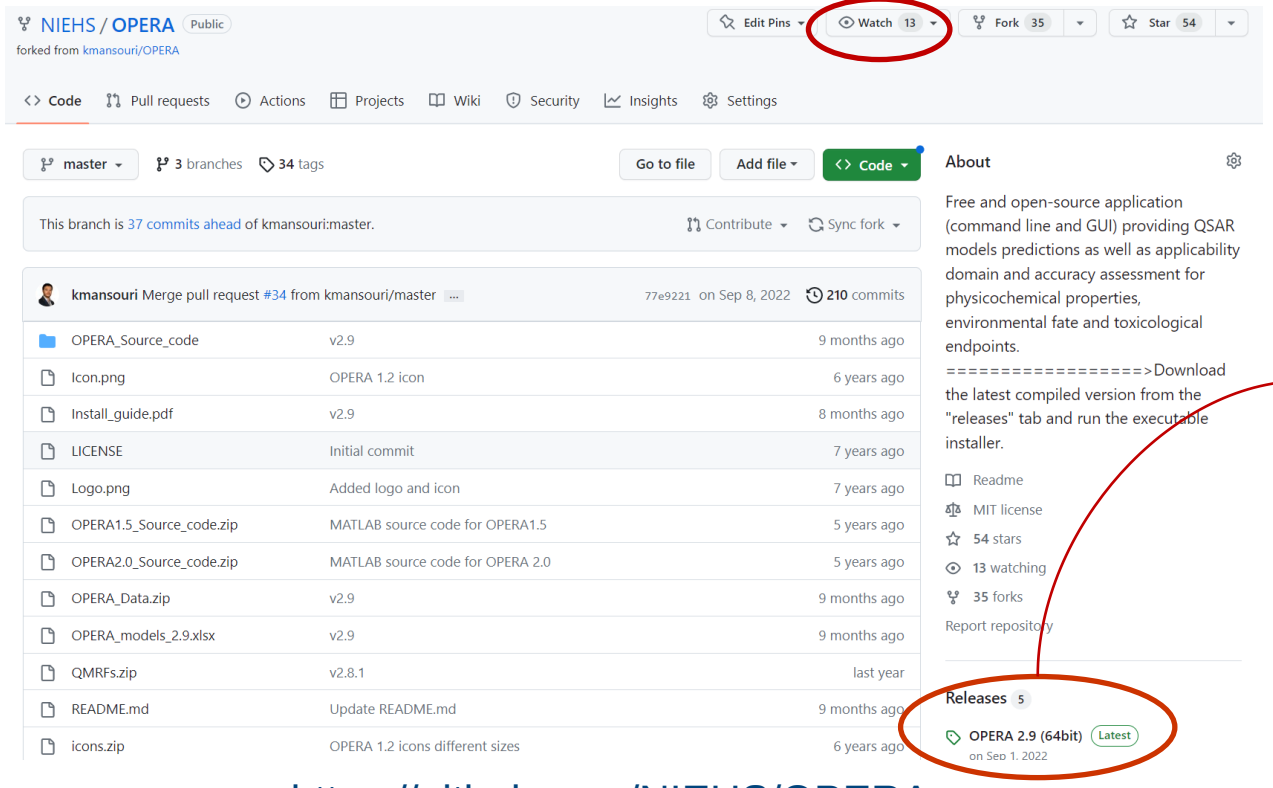
National Institute of Environmental Health Sciences
Division of Translational Toxicology

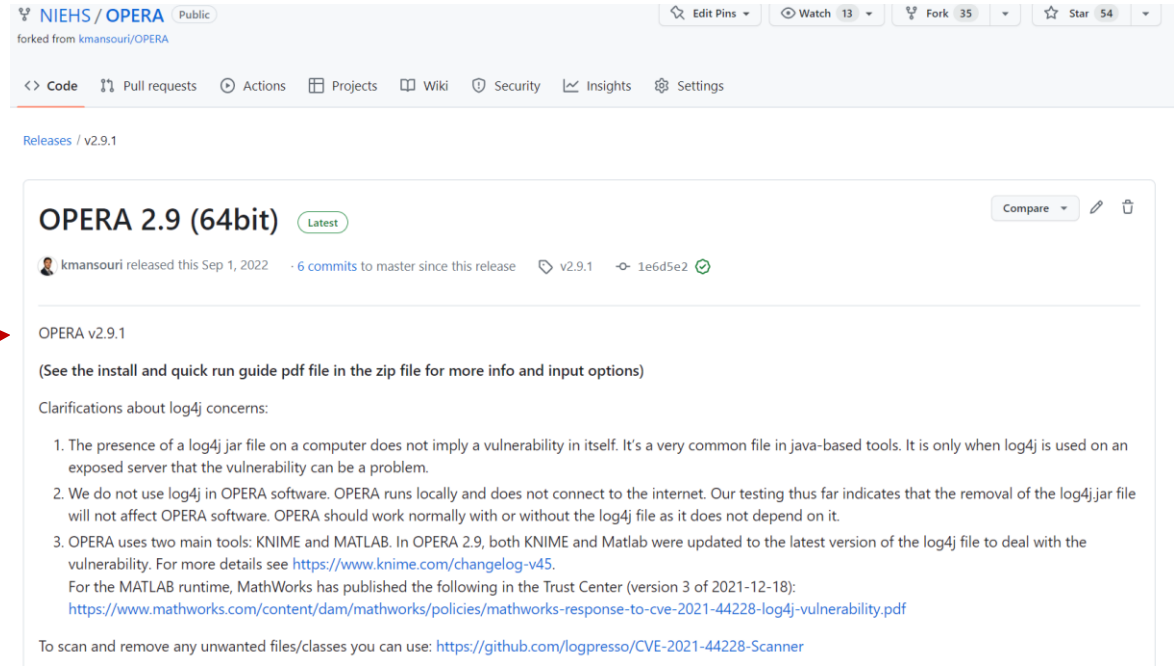
OPERA on GitHub

Source code

Packaged installers

Updates notifications





https://github.com/NIEHS/OPERA

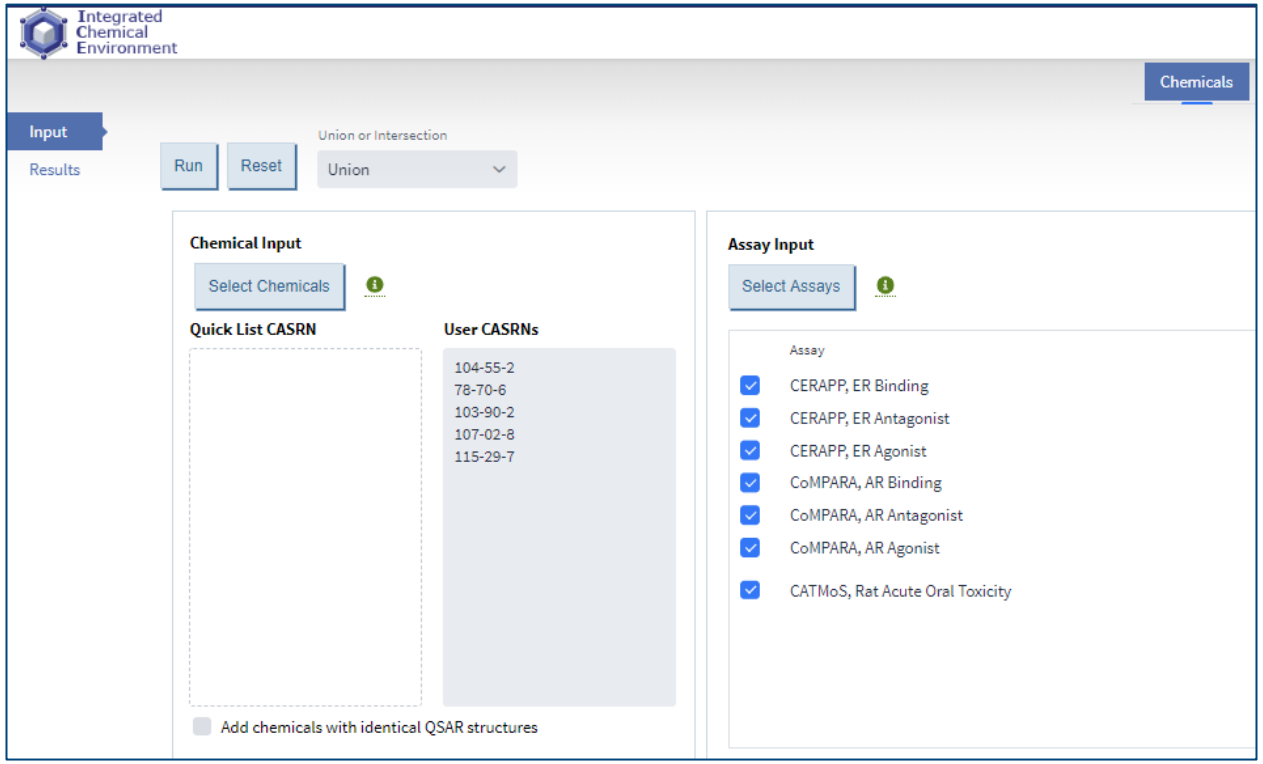
https://github.com/NIEHS/OPERA/releases

Over 8000 downloads

(https://tooomm.github.io/github-release-stats/)

National Institute of Environmental Health Sciences
Division of Translational Toxicology

Predictions on NTP/ICE



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

Download | Query Mixtures | Clear Filter | Number of chemicals = 5

Substance Name	CASRN	DTXSID	CATMoS, Rat Acute Oral Toxicity LD50	CoMPARA, AR Agonist Call	CoMPARA, AR Antagonist Call	CoMPARA, AR Binding Call	CERAPP, ER Agonist Call	CERAPP, ER Antagonist Call	CERAPP, ER Binding Call
Acetaminophen	103-90-2	DTXSID2020006	1625	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Endosulfan	115-29-7	DTXSID1020560	2.26	Inactive	Inactive	Inactive	Inactive	Active	Active
3-Phenylprop-2-enal	104-55-2	DTXSID1024835	2568	Inactive	Inactive	Inactive	Inactive	Active	Active
Acrolein	107-02-8	DTXSID5020023	40	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Linalool	78-70-6	DTXSID7025502	2097	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive

National Institute of Environmental Health Sciences
Division of Translational Toxicology

Demo

YouTube

Kamel.mansouri@nih.gov

National Institute of Environmental Health Sciences
Division of Translational Toxicology

Acknowledgments

The NICEATM Group (2024)

ICCVAM (ATWG & EcoWG)
EPA EFED
All international collaborators

Subscribe to NICEATM News email list
https://ntp.niehs.nih.gov/go/niceatm

National Institute of Environmental Health Sciences
Division of Translational Toxicology

Thank you for your attention!