Thursday, September 17, 2026

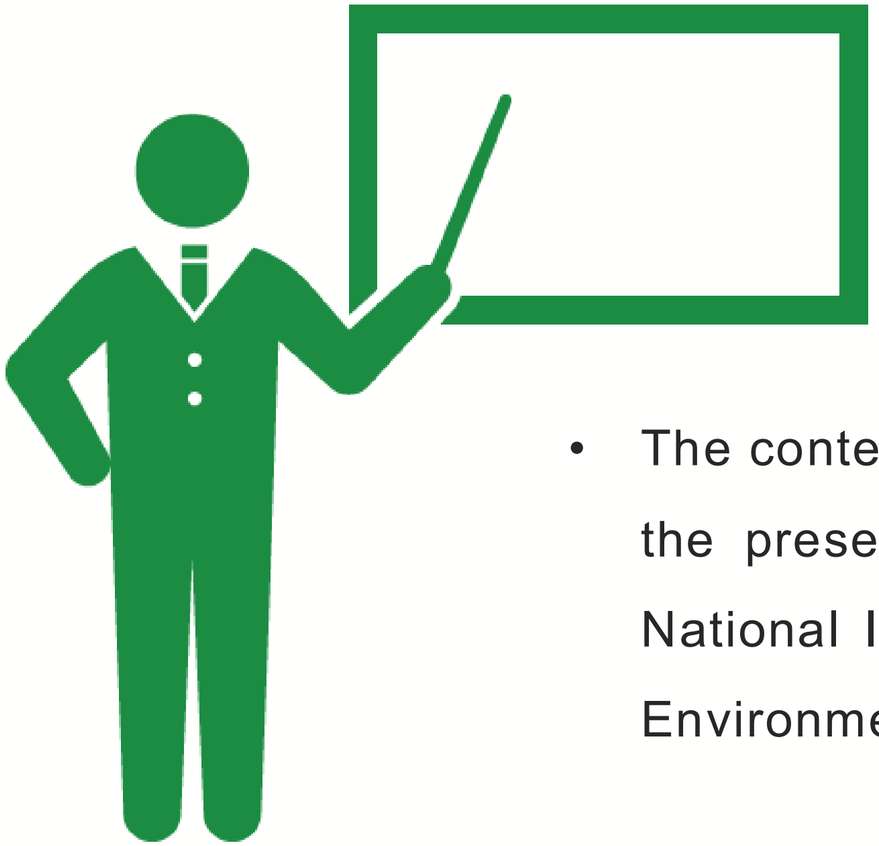
Eco-NAMs 4th webinar: QSARs to predict aquatic toxicity

KATE: an aquatic toxicity prediction system using QSAR

Director, Center for Environmental Regulatory Science (CERS),
Health and Environmental Risk Division,
National Institute for Environmental Studies (NIES), Japan

Koichi Ohno

National Institute for Environmenta Studies, Japar.



Disclaimers

- The contents of this presentation represent the personal views of the presenter and do not necessarily reflect the views of the National Institute for Environmental Studies, the Ministry of the Environment, Japan, or any other organizations.

- COI disclosure: The presenter has no conflicts of interest to disclose in relation to this presentation.

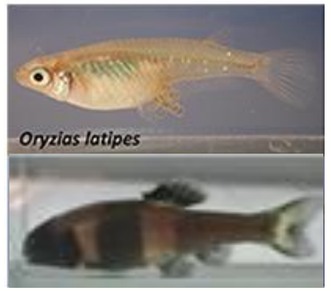
National Institute for Environmental Studies, Japan

Why ecotoxicity QSAR / NAMs?

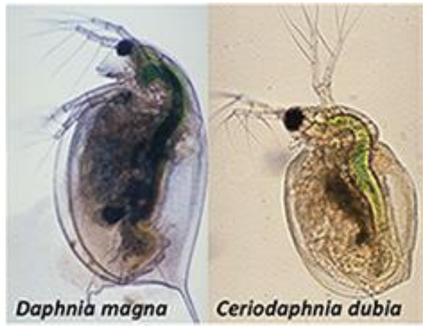
Regulatory ecotoxicity assessment generally requires hazard information for aquatic organisms.

However, measured data are not available for all chemicals and endpoints.

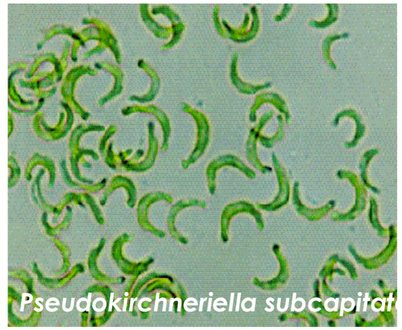
→ QSAR and read-across can help fill these gaps.



Fish



Daphnids (crustaceans)



Algae

National Institute for Environmental Studies, Japan

What is KATE2025?

- KATE (KAshinhou Tool for Ecotoxicity)
  (Kashinhou: CSCL, Chemical Substances Control Law in Japan)
  https://kate.nies.go.jp/

- An ecotoxicity QSAR system developed by the Health and Environmental Risk Division of the National Institute for Environmental Studies (NIES) under contract with the Ministry of the Environment, Japan.

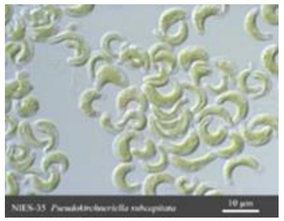
- The latest version is KATE2025 version 2.0, released on March 30, 2026.
  The first version, KATE2011, was followed by major updates in 2017, 2020, 2025.

- The reference substance data used to build KATE consist of ecotoxicity test results* conducted by the Ministry of the Environment (fish, daphnids, and algae) and fish acute toxicity test results from the U.S. EPA fathead minnow database.

- The tests conducted by the Ministry of the Environment were performed in GLP-compliant test facilities using methods conforming to OECD Test Guidelines.
  Cf. results of aquatic toxicity tests of chemicals conducted by MOE (as of March 2025).
  https://www.env.go.jp/content/000316616.pdf

National Institute for Environmental Studies, Japan

Ecotoxicity endpoints predicted by KATE2025



Predicted Toxicity Type		Species (Scientific Name)	Testing Method	Duration	Toxicity endpoint
Organisms	Acute/ Chronic				
Fish	Acute	*Oryzias latipes*, Medaka *Pimephales promelas*[*1] Fathead minnow	Fish acute toxicity test (OECD TG 203) [7]	96 h	LC50
Daphnid	Acute	*Daphnia magna*	Daphnia magna immobilization test (OECD TG 202) [9]	48 h	EC50
Alga	Acute	*Raphidocelis subcapitata* [*2]	Algal growth inhibition test (OECD TG 201) [11]	72 h	EC50
Fish	Chronic	*Oryzias latipes*	Fish early-life-stage toxicity test (OECD TG 210) [8]	Embryonic stage and 30 days after hatching [*3]	NOEC
Daphnid	Chronic	*Daphnia magna*	Daphnia magna reproduction test (OECD TG 211) [10]	21 d	NOEC
Alga	Chronic	*Raphidocelis subcapitata* [*2]	Algal growth inhibition test (OECD TG 201) [11]	72 h	NOEC

Ref) KATE2025 Technical Document: https://kate.nies.go.jp/doc/KATE2025_tech_doc-en.pdf

*1 Medaka data are prioritized when both species are available.

*2 Former name: *Pseudokirchneriella subcapitata*

*3 Test duration varies depending on fish species and hatching time.

National Institute for Environmental Studies, Japan

Number of training data (n) and chemical classes (as of KATE2025ver1.1)

	n	#QSARs that satisfy R2≧0.7, Q2≧0.5, n≧5	#QSARs do not satisfy the criteria*
Fish Acute	1079	51	70
Daphnid Acute	486	28	77
Alga Acute	554	11	69
Fish Chronic	33	3	7
Daphnid Chronic	384	12	77
Alga Chronic	534	16	83

* In KATE, QSAR classes (*i.e.* endpoint-specific QSAR models) with excess toxicity, such as reactive classes, are less QUANTITATIVELY predictive than those of unreactive chemical classes.

National Institute for Environmental Studies, Japan

Recent Activity: Test Data Quality Recheck

- Training set data
  - MOE's aquatic toxicity testing results and US EPA's Fathead minnow database (only for fish acute toxicity)
  - MOE data has been reviewed since 2021 to check consistency with revised OECD Test Guidelines and Guidance Document No. 23.
    - ✓Toxicity values were recalculated based on the latest guidance document, if necessary.
    - ✓Several chemicals' data were excluded from regression equation e.g., when the test chemical was a mixture, when the purity of the chemical was low.
  - KATE was updated to KATE2025 after finishing the review of aquatic toxicity test for all six toxicity endpoints.

How KATE predicts ecotoxicity

- Structural class classification
- Log-log regression

National Institute for Environmental Studies, Japan

National Institute for Environmental Studies, Japan

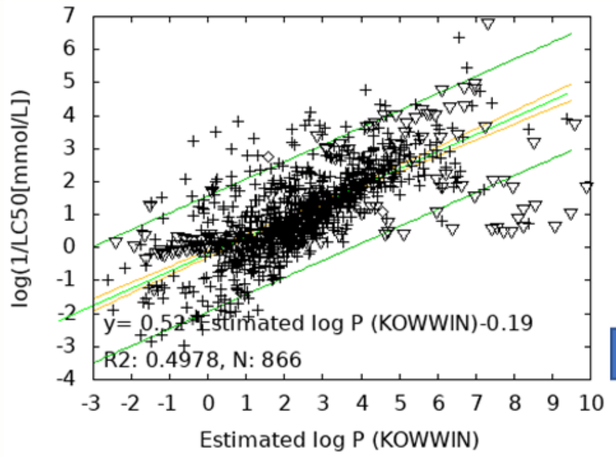
KATE predicts ecotoxicity based on chemical structural classification and linear regression: log(1/toxicity) vs logP

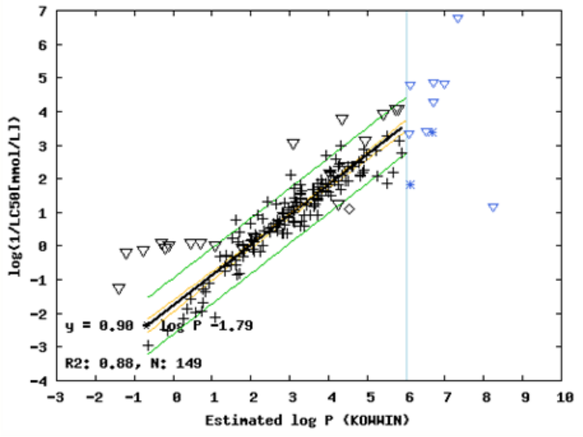
The y-axis is log(1/toxicity [mmol/L]), so higher positions indicate stronger toxicity (smaller toxicity values).

Narcotic class (baseline toxicity)

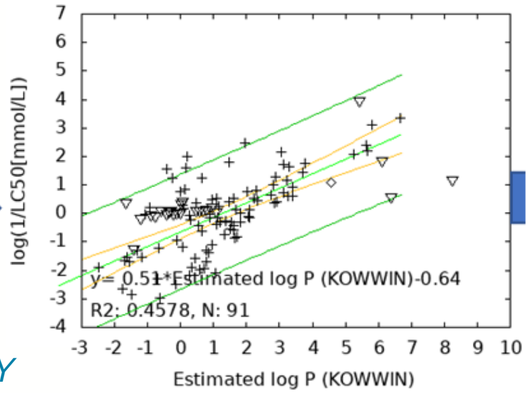
Condition X

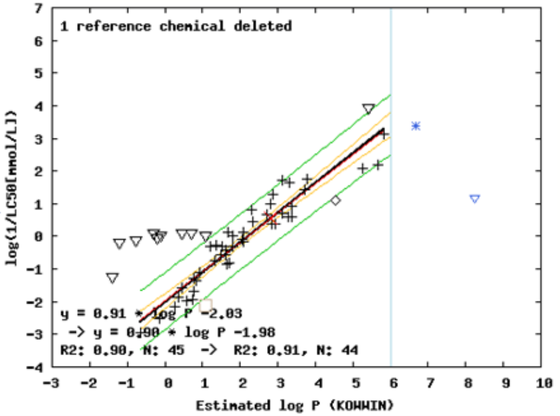
Whole available data of fish acute test





CO_X alcohol unreactive Fish Acute Class





Condition Y

Ex. Alcohol

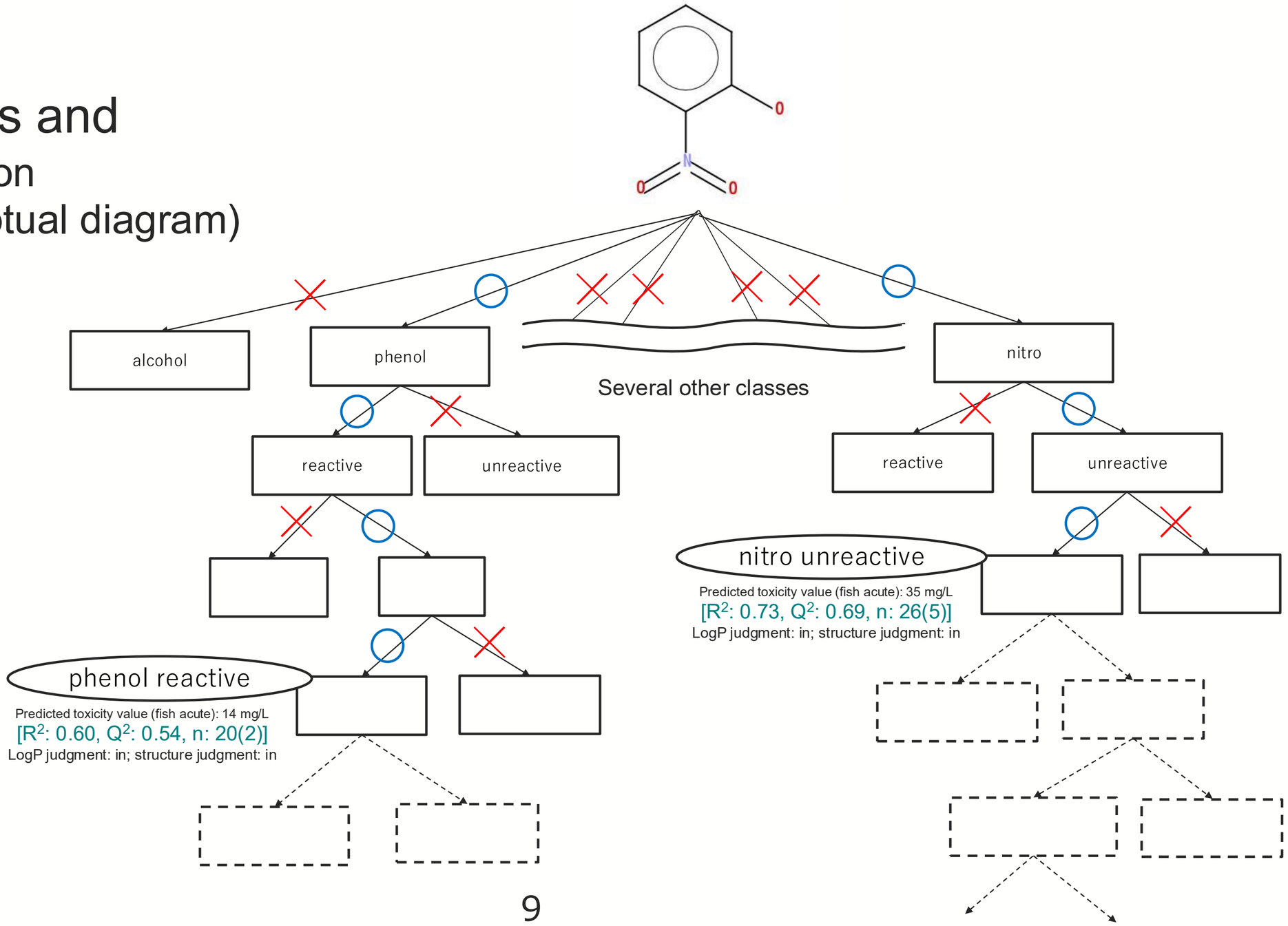
Condition Y and A and B

Ex. A= CO_X, B=unreactive

* KATE creates separate QSAR regression equations for different structural classes to predict ecotoxicity.

National Institute for Environmental Studies, Japan

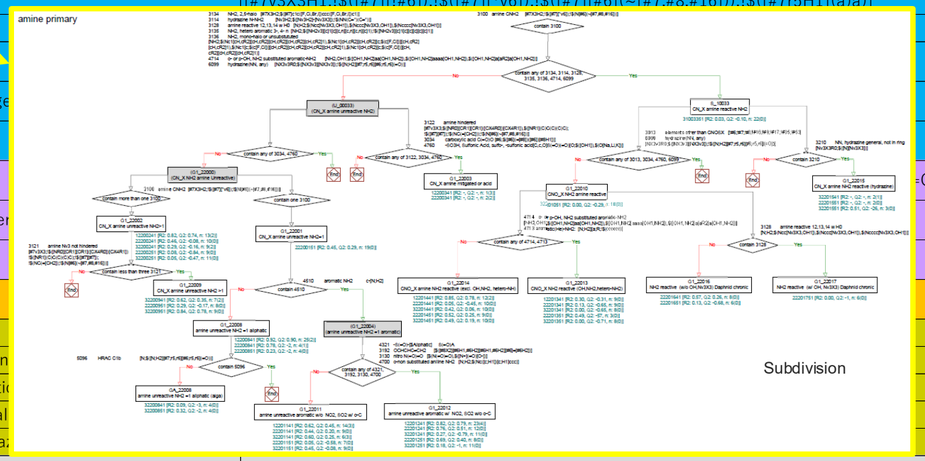
Structural-class and QSAR classification by KATE (conceptual diagram)



National Institute for Environmental Studies, Japan

Major categories of KATE chemical classes

	Major Class	ID	Explanation of the class	SMARTS
1	**acid**	3034	carboxylic acid C(=O)O	[#6;$([#6](=[#8])([#6])[#8H1])]
		4760	-SO3H, Sulfonic Acid, sulfo-, -sulfonic acid	[C,c,O]S(=O)(=O)[O;$([OH1]),$(O[Na,Li,K])]
2	**alcohol**	3046	alcohol COH	[#6;$([#6][#8;H]);!X3;v4]
3	**aldehyde, ketone**	3031	ketone CC(=O)C	[#6;$([#6](=[#8])([#6])[#6])]
		3036	aldehyde	[#6H1;$([#6](=[#8])[#6])]
4	**ester**	3032	ester CC(=O)OC	[#6;$([#6](=[#8])([#6])[#8][#6]);!$([#6](=[#8])([#6])[#8][#6]=[O,S,N])]
		3145	acetal	[#6X4;$([#6]([#8])[#8])]
5	**ether**	3044	ether general	[#8H0;!$([#8]C=[O,N,S]);!$([#8]C[#8]);$([#8]([#6])[#6])]
6	**phenol**	3047	phenol cOH	[OX2H][cX3;$(c1ccccc1)]
7	**amine primary**	3100	amine CNH2	[#7X3H2;!$([#7][*v6]);!$(N[#6](~[#7,#8,#16]))]
8	**amine sec, tert**	3110	amine CNH1	[#7v3X3H1;!$([#7][#6]);!$([#7][*v6]);!$([#7][#6](~[#7,#8,#16]));!$([#7r5H1](a)a)]
		3120	amine CNH0	
9	**aromatic n**	4911	aromatic n	
10	**hydrazine**	3210	NN, hydrazine ge	
11	**nitrile**	3104	nitril C#N	
12	**amide**	3123	amide	
13	**carbamate**	3041	carbamate gener	
14	**nitro**	3130	nitro N(=O)=O	
15	**phosphorus**	5018	Phosphorus [P]	
16	**sulfur**	3234	SH thiophenol	
		3235	SH mercaptan, n	
		3535	disulfide excepti	
		3536	disulfide general	
		3537	2-mercapto-thia;	
17	**halogen**	4507	halogen	[F,Cl,Br,I]
18	**heteroatomatic**	4911	aromatic n	[n]
		4912	aromatic o	[o]
		4913	aromatic s	[s]
	Others	3106	azo N=N	[NX2;$(N=N)]
		4541	epoxide monocyclic	[#8r3;$([#8]1[#6R1][#6R1])]
		3108	imino C=N-, guanidine	[#7v3X2;$(N=[Cv4X3]);!$(N[N,O,S])]
			etc	
			.	
			.	
			.	



Containing oxygen

Containing nitrogen

Containing oxygen and nitrogen

Containing phosphorus

Containing sulfur

Other

There is also a major class of narcotic substances (baseline toxicity class).

Substances not classified into any class are assigned to the "Unclassified" class and are not displayed by default.

Example

- Toxicity prediction
- Statistical criteria (R2≧0.7, Q2≧0.5, n≧5)
- Detail information (Verify QSAR)
- Applicability Domain Check

National Institute for Environmental Studies, Japan

National Institute for Environmental Studies, Japan

Entry page for KATE2025 (https://kate.nies.go.jp/)

■ Main Content ■ Site Map ■ Japanese Page

KAshinhou Tool for Ecotoxicity
Ecotoxicity prediction system

国立研究開発法人 国立環境研究所 National Institute for Environmental Studies

KATE2025 | KATE on PAS 2011 Japanese | ALDAME (working title) *β version* | Change Log | Site Policy | FAQ

HOME > *KATE2025*

KATE2025

KATE2025 login

◆ How to use

The instructions are provided in the Operating manual. KATE supports the browser Firefox.

- KATE2025 login (version 2.0)
- KATE2025 Operating manual (PDF 7.8MB)
- KATE2025 Technical Document (PDF 9.3MB)
- QMRF (PDF 0.8MB)

◆ About KATE2025

KATE2025 is an major updated version of KATE2020.
For the main changes from "KATE2020", please see the Change Log.

■ Features

- Prediction of toxicity values for an input chemical

(Acute)

- 50% lethal concentration (LC_{50}) in the fish acute toxicity test (OECD TG 203)
- 50% effective concentration (EC_{50}) in the *Daphnia magna* acute immobilization test (OECD TG 202)
- 50% effective concentration (EC_{50}) in the algal growth inhibition test (OECD TG 201)

(Chronic)

- No-observed-effect concentration (NOEC) in the fish early-life-stage toxicity test (OECD TG 210)
- No-observed-effect concentration (NOEC) in the *Daphnia magna* reproduction test (OECD TG 211)
- No-observed-effect concentration (NOEC) in the algal growth inhibition test (OECD TG 201)

- Judgment whether the prediction results are within the applicability domain for structure and log P
- Display of QSAR model graph
- Prediction of toxicity values for multiple chemicals

KAshinhou Tool for Ecotoxicity
KATE2025 version 2.0

Terms of Agreement

```
KOWWIN v1.69 (April 2015)

© 2000-2015 U.S. Environmental Protection Agency

KOWWIN is owned by the U.S. Environmental Protection
Agency and is protected by copyright throughout the
world.

Permission is granted for individuals to download and
use the software on their personal and business
computers.

Users may not alter, modify, merge, adapt or prepare
derivative works from the software. Users may not
remove or obscure copyright, tradename, or proprietary
notices on the program or related documentation.

KOWWIN contained therein is a tradename owned by the
U.S. Environmental Protection Agency.
```

1) Check the box

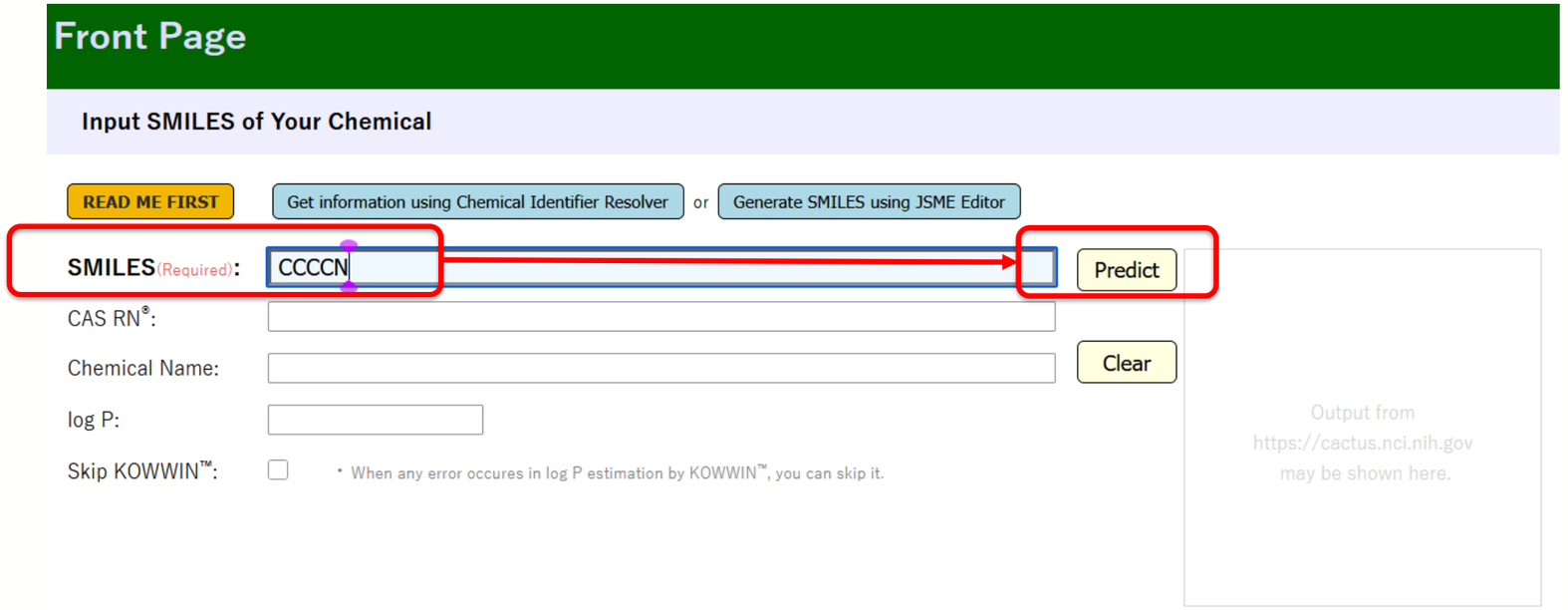
☐ I agree to and accept the terms of agreement above.

Enter

2) Click

National Institute for Environmental Studies, Japan

KATE2025: Input page



▼ *Help:* ***SMILES notation in KATE2025***

SMILES (Simplified Molecular Input Line Entry System) is a line notation that expresses molecular structural formulas as strings. There is a detailed explanation on Daylight's web page.

KATE2025 can predict ecotoxicity of only organic chemicals except for some inorganic nitrogen compounds such as hydrazine.

KATE2025 cannot predict ecotoxicity of chemicals represented as following types of SMILES:

i. SMILES which does not contain carbon or nitrogen atoms.
ii. SMILES which includes elements other than H, C, N, O, F, Si, P, S, Cl, As, Br, Sn, and I.
iii. SMILES which includes ions other than ammonium [N+] or [n+].
iv. SMILES which includes ".", i.e. SMILES which expresses a mixture.

The strings such as [Na], [K], [Li], [Na+], [K+] and [Li+] in SMILES should be replaced by the protonated forms. For example, SMILES "c1ccccc1O[Na]" needs to be replaced by "c1ccccc1O".

SMILES can be converted from two-dimensional structural formula using chemical software such as JMSE.

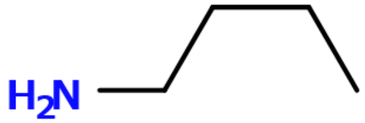
Results

Case of n-butylamine (CCCCN)

Summary of the Query Chemical

Query Chemical

SMILES	CCCCN	
CAS RN®	109-73-9	
Chemical Name	n-Butylamine	
log P	User Input Value	Re-calculate
	Estimated Value by KOWWIN™	0.83
	Measured Value in KOWWIN™ Database	0.97
Molecular Weight	73.14	



In this case, only the fish acute and daphnid acute QSARs satisfied statistical criteria.

QSAR Prediction Result

Toxicity filter:

	all	Fish	Daphnid	Alga
Acute	☑	☑	☑	☑
Chronic	☑	☑	☑	☑

Statistical filter:

☐ Apply the following statistical criteria (uncheck this box to show all QSAR results):
☐ $R^2 \geq$ 0.7 ☐ $Q^2 \geq$ 0.5 ☐ n ≥ 5

Click!

Print Detail	QSAR Class Name*2 (Click the class name to see the QSAR details)	Type of Predicted Toxicity*3: Organism	Type of Predicted Toxicity*3: Acute or Chronic	Predicted Toxicity [mg/L]	95% Prediction Interval	Applicability Domain Judgement: log P*4 [Range]		Applicability Domain Judgement: Structure*5	Statistics of QSAR Regression: R^2	Q^2 *6	RMSE	n*7	statsical criteria*8
☑	amine primary unreactive NH2=1 aliphatic	Fish	Acute	84	[5.4, 1300]	in	[-1.61, 5.25]	in	0.84	0.81	0.54	26(2)	✓
☑	amine primary unreactive NH2=1 aliphatic	Daphnid	Acute	16	[0.56, 490]	in	[-1.61, 4.76]	in	0.82	0.66	0.51	9(1)	✓
☐	amine primary unreactive NH2=1 aliphatic alga	Alga	Acute	3.0	[0.0050, 1800]	in	[-1.61, 4.76]	in	0.55	0.21	0.98	9(0)	
☐	CNO_X unreactive Fish Chronic, w/ N,O	Fish	Chronic	0.18	[0.0079, 4.2]	in	[-1.61, 5.99]	in	0.62	0.54	0.57	19(2)	
☐	amine primary unreactive NH2=1 aliphatic	Daphnid	Chronic	0.89	[0.015, 54]	in	[-1.61, 1.63]	in	0.23	-2.22	0.26	4(0)	
☐	amine primary unreactive NH2=1 aliphatic alga	Alga	Chronic	0.47	[0.0022, 98]	in	[-1.61, 4.76]	in	0.66	0.41	0.82	9(0)	

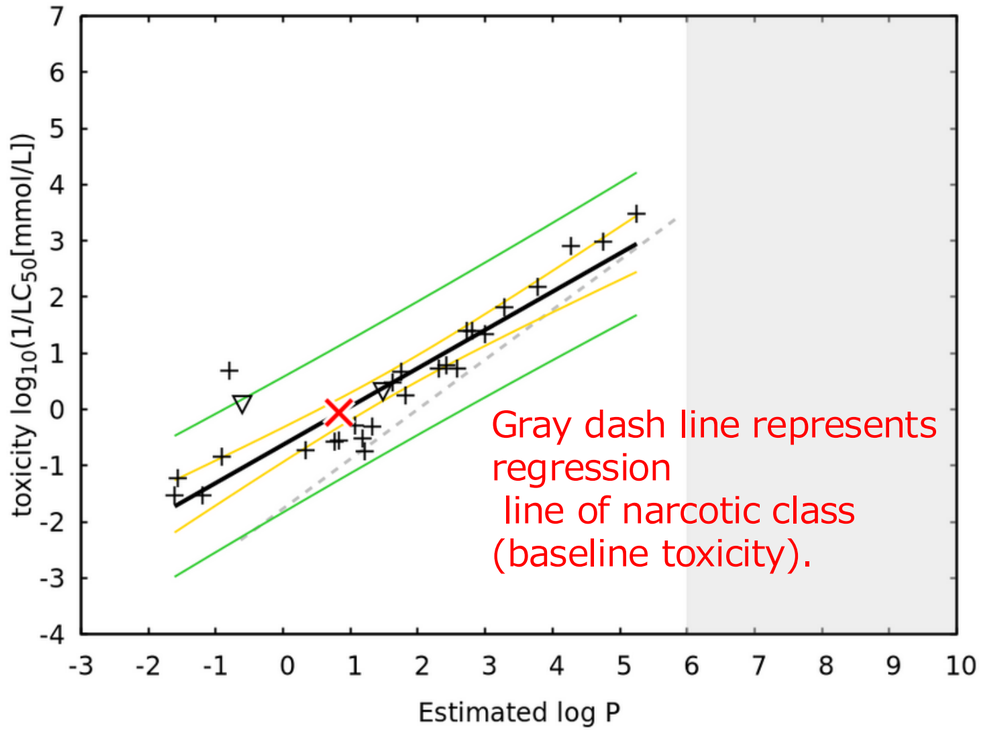
Create Print Format

Verify QSAR

Type: **Fish (Acute)** Structure Class ID: **G1_22008** QSAR Class Name: **amine primary unreactive NH2=1 aliphatic**

Regression Equation

Horizontal Axis: log P | Go | Default

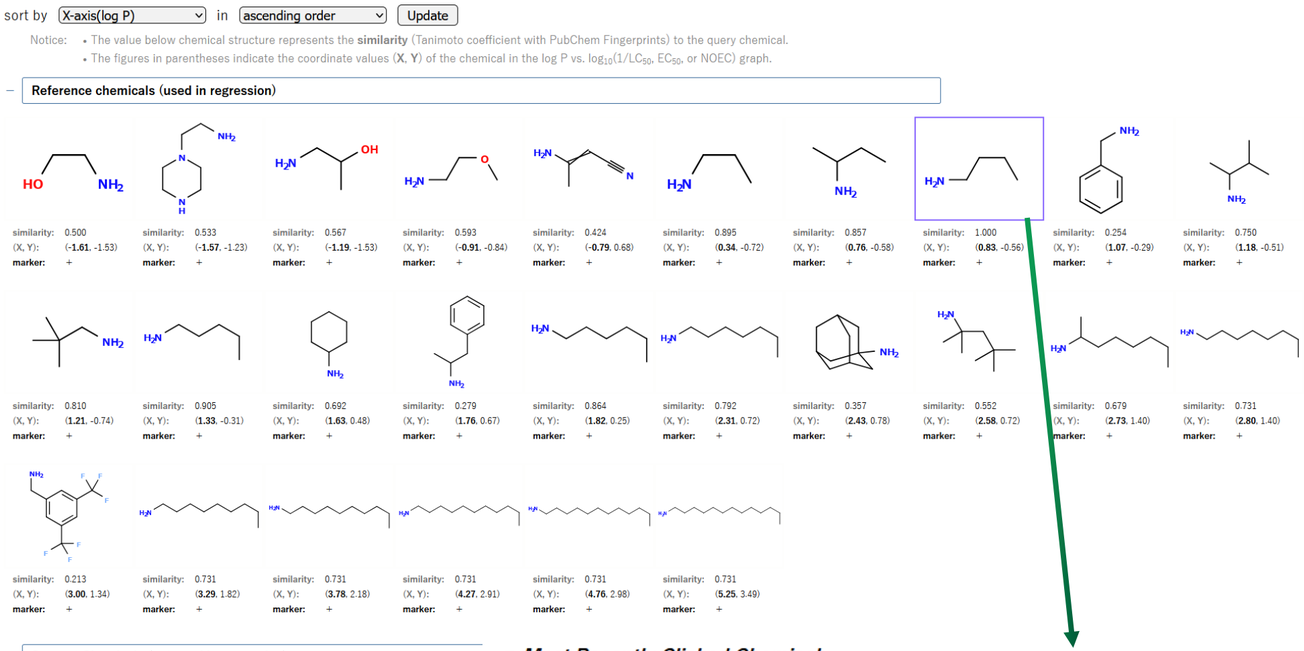


Gray dash line represents regression line of narcotic class (baseline toxicity).

Exclude the data other than (+)

Equation	X (Estimated log P)	Y ($log_{10}(1/LC_{50}$ [mmol/L]))	Predicted Toxicity [mg/L]	95% Prediction Interval	Applicability Domain Judgement			Statistics of QSAR Regression				
					log P [Range]		Structure	R^2	Q^2	RMSE	n	statsical criteria
Y = 0.68 * X - 0.63	0.83	-0.06	84	[5.4, 1300]	in	[-1.61, 5.25]	in	0.84	0.81	0.54	26(2)	✓

By clicking QSAR name, users can view detailed information of the QSAR.



▼ ***Most Recently Clicked Chemical***

SMILES: CCCCN
CAS RN®: 109-73-9
Chemical Name: Butylamine
Molecular Weight: 73.14
(X, Y): (0.83, -0.56)
Square of Residual: 0.25
Measured Toxicity Value Data:
LC_{50} [mg/L]: 268.0, Species: *Pimephales promelas*, Reference: USEPA
Note:

In this example, the same chemical is found in the test dataset.
Measured toxicity 268.0 mg/L
Predicted toxicity 84 mg/L

National Institute for Environmental Studies, Japan

Applicability Domain Judgement

QSAR Prediction Result

Toxicity filter:

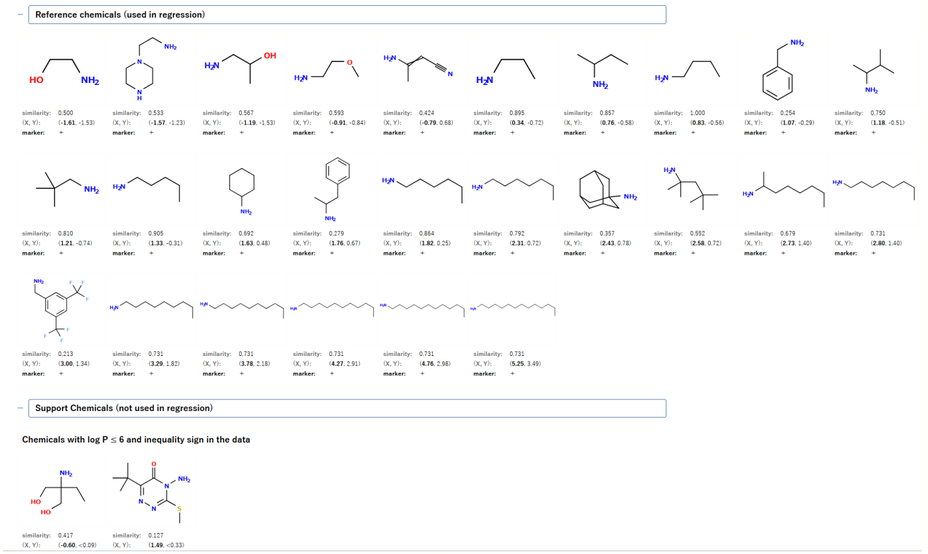
	all	Fish	Daphnid	Alga
Acute	☑	☑	☑	☑
Chronic	☑	☑	☑	☑

Statistical filter:

☑ Apply the following statistical criteria (uncheck this box to show all QSAR results):

☑ $R^2 \geq$ 0.7 ☑ $Q^2 \geq$ 0.5 ☑ $n \geq$ 5

Print Detail	QSAR Class Name*2 Click the class name to see the QSAR details	Type of Predicted Toxicity*3		Predicted Toxicity [mg/L]	95% Prediction Interval	Applicability Domain Judgement			Statistics of QSAR Regression				
		Organism	Acute or Chronic			log P*4 [Range]		Structure*5	R^2	Q^2 *6	RMSE	n*7	statsical criteria*8
☑	amine primary unreactive NH2=1 aliphatic	Fish	Acute	84	[5.4, 1300]	in	[-1.61, 5.25]	in	0.84	0.81	0.54	26(2)	✓
☑	amine primary unreactive NH2=1 aliphatic	Daphnid	Acute	16	[0.56, 490]	in	[-1.61, 4.76]	in	0.82	0.66	0.51	9(1)	✓



(1) log P

logP of the target chemical should be within the min and max values of test chemicals.

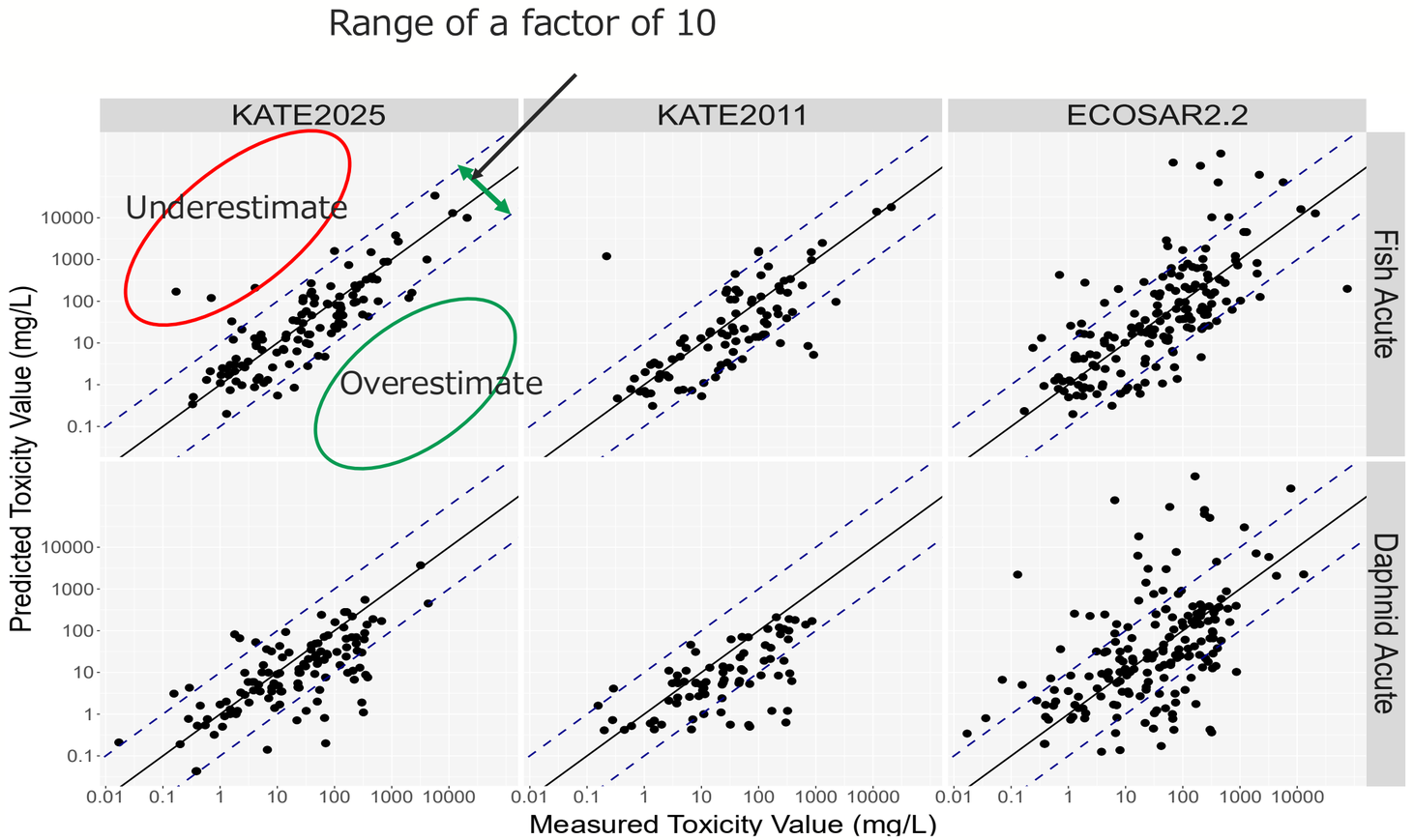
(2) Structural applicability domain (in KATE)

A query chemical is judged **"in"** if its relevant structural features are represented among the chemicals used to build the QSAR regression; otherwise, it is judged **"out."**

National Institute for Environmental Studies, Japan

Predictive Performance of KATE2025

Experimental data on fish and daphnid acute toxicity of OECD high production volume (HPV) chemicals were collected as validation data set.
(Fish Acute: 178, Daphnid Acute: 196)



		KATE2025	KATE2011	ECOSAR2.2
1. How many chemicals predicted within model's AD	Fish Acute	114(64%)	90(51%)	167(94%)
	Daphnid Acute	117(60%)	82(42%)	182(93%)
2. (for chemicals predicted within AD) How many predictions fell within a factor of 10 to experimental toxicity value	Fish Acute	102(89%)	78(87%)	127(76%)
	Daphnid Acute	96(82%)	65(79%)	96(67%)

National Institute for Environmental Studies, Japan

Pros and Cons of the Current KATE System

Pros:

- Training dataset in KATE is transparent and reliable.
  Only MOEJ and USEPA(fish) data are used.
- The classification process is transparent and interpretable.
- The applicability domain, including statistical criteria, is narrower than that of such as ECOSAR; however, predictive performance is generally higher for chemicals within the domain.
- KATE uses a structural applicability domain to reduce the risk of underestimating toxicity.

National Institute for Environmental Studies, Japan

Pros and Cons of the Current KATE System

Cons:

- Chemical classes that has some specific (i.e., excess) toxicity cannot be predicted quantitatively.
  - However, qualitative evaluation may still be feasible depending on the chemical class.

- Classification is based solely on chemical structure. While structure similar chemicals tend to exhibit similar ecotoxicity, KATE does not currently provide mechanistic explanations of toxicity. The classification scheme is relatively complex and partly heuristic, and further refinement is needed.
  - Therefore, it is recommended to combine KATE with other approaches, such as structural alerts or mechanistic models.
  - Also, for better practice, QSARs including KATE should be used in combination with other approaches, such as read-across methods.

National Institute for Environmental Studies, Japan

Current regulatory use of KATE in Japan

- Since KATE is named after “Kashinhou” (CSCL; Chemical Substance Control Law), its main development objective, is regulatory risk assessment in Japan.

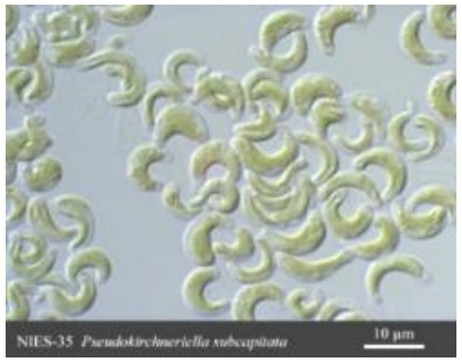
- But, at present, QSAR-estimated toxicity values are not used to derive regulatory or assessment values. Alternative toxicity values are not adopted as official regulatory values.

- Under CSCL, KATE is used in the screening assessments as a reference tool (i.e., not directly used to derive Assessment Values).

- As for Initial Risk Assessment of Environmental Chemicals performed by MOE, KATE is used both for QSAR and for Read Across methods. The estimated toxicity values derived from QSAR and read-across are used to support expert judgement.
  https://www.env.go.jp/en/chemi/chemicals/profile_erac/

National Institute for Environmental Studies, Japan

Thank you for listening!



ohno.koichi@nies.go.jp

National
Institute for
Environmental
Studies, Japan

supplement

National Institute for Environmental Studies, Japan

Number of training data (n) and chemical classes (as of KATE2025ver1.1)

	n	#QSARs that satisfy R2≧0.7, Q2≧0.5, n≧5	#QSARs do not satisfy the criteria*
Fish Acute	1079	51	70
Daphnid Acute	486	28	77
Alga Acute	554	11	69
Fish Chronic	33	3	7
Daphnid Chronic	384	12	77
Alga Chronic	534	16	83

* In KATE, QSAR classes (*i.e.* endpoint-specific QSAR models) with excess toxicity, such as reactive classes, are less QUANTITATIVELY predictive than those of unreactive chemical classes.

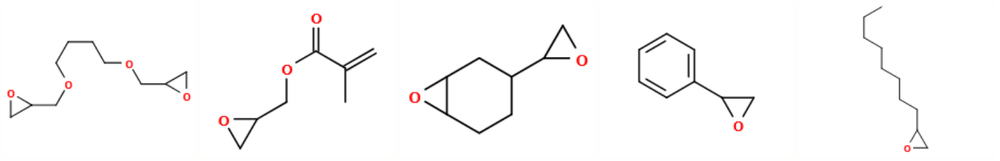
National Institute for Environmental Studies, Japan

Verify QSAR

Type: **Fish (Acute)** Structure Class ID: **G1_21046** QSAR Class Name: **CO_X epoxide**

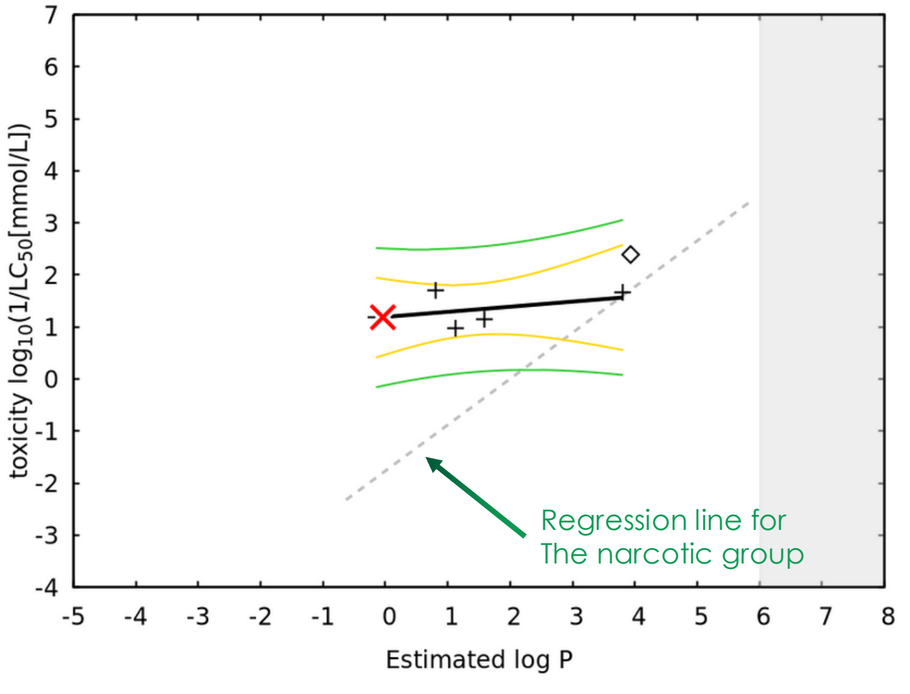
Regression Equation

Reference chemicals (used in regression)



similarity:	0.433	0.317	0.289	0.171	0.448
(X, Y):	(**-0.15**, 1.19)	(**0.81**, 1.71)	(**1.13**, 0.97)	(**1.59**, 1.14)	(**3.81**, 1.66)
marker:	+	+	+	+	+

Horizontal Axis: log P | Go | Default



Query Chemical

SMILES	C1CO1
CAS RN®	
Chemical Name	
log P (Estimated)	-0.05
Molecular Weight	44.05

Plot Legends

Query Chemical — Regression

Click to show the plot legends

Reference Chemicals (used in regression)

+ : log P ≤ 6.0 (no "<" or ">")

Plot Controls

XY: − + X: − + Y: − +

Exclude the data other than (+)

Equation	X (Estimated log P)	Y ($\log_{10}(1/LC_{50}$ [mmol/L]))	Predicted Toxicity [mg/L]	95% Prediction Interval	Applicability Domain Judgement			Statistics of QSAR Regression				
					log P [Range]		Structure	R^2	Q^2	RMSE	n	statsical criteria
Y = 0.10 * X + 1.19	-0.05	1.19	2.9	[0.14, 60]	in	[-0.15, 3.81]	in	0.19	-1.22	0.27	5(1)	

Some specific toxicity does not correlate with LogP and thus we cannot predict it QUANTITATIVELY.
But this class could be predicted QUALITATIVELY and might be used as SAR.
KATE does not have a function as SAR at the current version, though.

QSAR classes that do not meet the statistical criteria may still provide useful information for qualitative toxicity assessment or read-across.