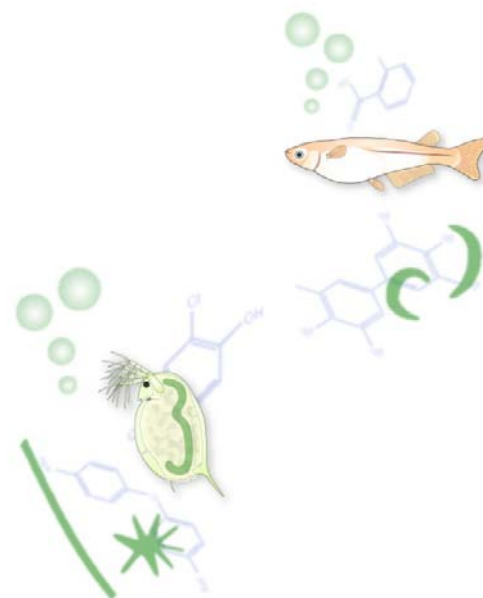




生態毒性QSARモデル 「KATE」について

 独立行政法人国立環境研究所

環境リスク研究センター 蓮沼 和夫

生態毒性QSARの状況



代表的な生態毒性QSAR

▶ 代表的なQSARの特徴

名称	開発元	記述子	予測エンドポイント	その他
KATE	環境省、 国立環境研究所	logP (水-オクタノール 分配係数)	魚類・甲殻類急性毒性 (魚類・甲殻類慢性、藻類は開発中)	ドメイン判定: 構造、 記述子
TIMES	ブルガリア ブルガス大学	logBCF _{tox} 、 LUMO等	魚類・甲殻類急性毒性等 (<i>Rana japonica</i> , <i>Lymnaea stagnalis</i> , <i>Carassius auratus</i> , <i>Oryzias latipes</i> , <i>Leuciscus idus</i> , <i>Pimephales promelas</i> , <i>Daphnia magna</i> , <i>Daphnia pulex</i> , <i>Ceriodaphnia dubia</i> , <i>Escherichia coli</i> , <i>Bacillus subtilis</i> , <i>Tetrahymena pyriformis</i> 等)	・ドメイン判定: 構造 ・有償(約150万/年)
ECOSAR	米国EPA	主にlogP	魚類・甲殻類・藻類急性毒性 魚類・甲殻類・藻類ChV (ChV: Chronic Value: NOECとLOECの幾何平均)	ドメイン判定: 記述子
OECD QSAR Toolbox	OECD、EU	任意(ユーザ がlogPやpKa 等から選択)	任意	・ドメイン判定: ユー ザが判断 ・1物質毎にユーザ がQSAR式を構築 する必要あり



代表的な生態毒性QSARの予測精度の現状

▶ 実測毒性値とその予測値を比較

- ▶ 三省合同審議会※に予測結果を提出しているKATE、TIMES、ECOSARを使用（KATEは、Looを実施）
- ▶ ドメイン外の予測結果は除外し検討

・KATE: logP判定 ×、構造C判定: ×

・ECOSAR: logP判定 ×

・TIMES: Reactive Unspecifiedクラス、構造ドメイン外

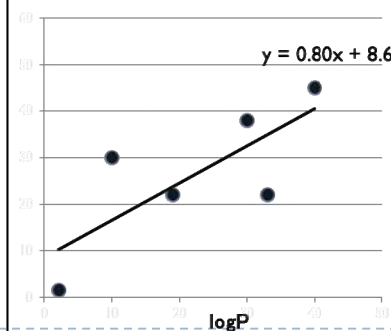
▶ 検証に用いた実測毒性

- ▶ KATE2011年3月版の
参照物質

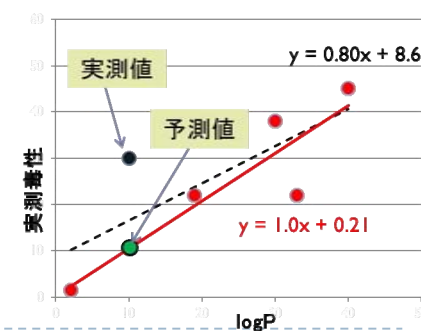
※: 薬事・食品衛生審議会薬事分科会化学物質安全対策部会化学物質調査会、化学物質審議会審査部会、中央環境審議会環境保健部会化学物質審査小委員会

【補足】 leave-one-out法について

- ▶ KATE: 参照物質の毒性を予測しても意味がない
- ▶ 実測値と予測値を比較する1物質を除いてQSAR式を作成し、その1物質の毒性値を予測 (leave-one-out)

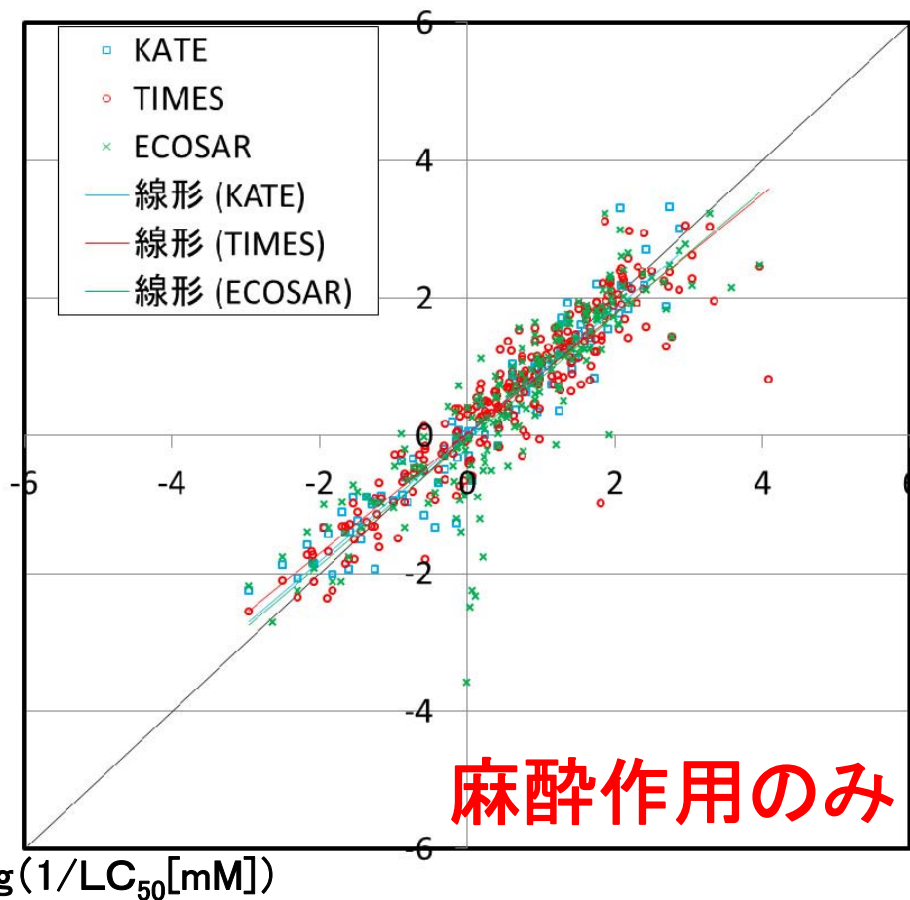
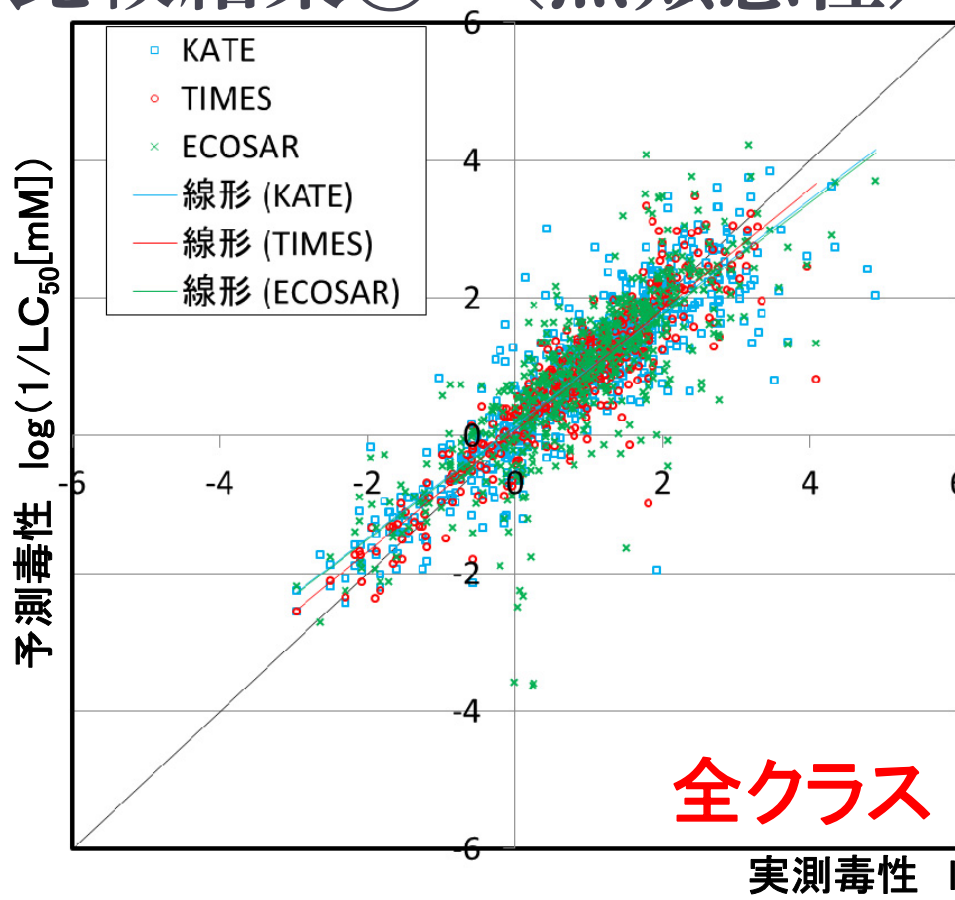


QSAR式
(実測毒性がある6物質より作成)



leave-one-out QSAR式
(1物質除いた5物質(赤丸)より作成)

比較結果① (魚類急性)



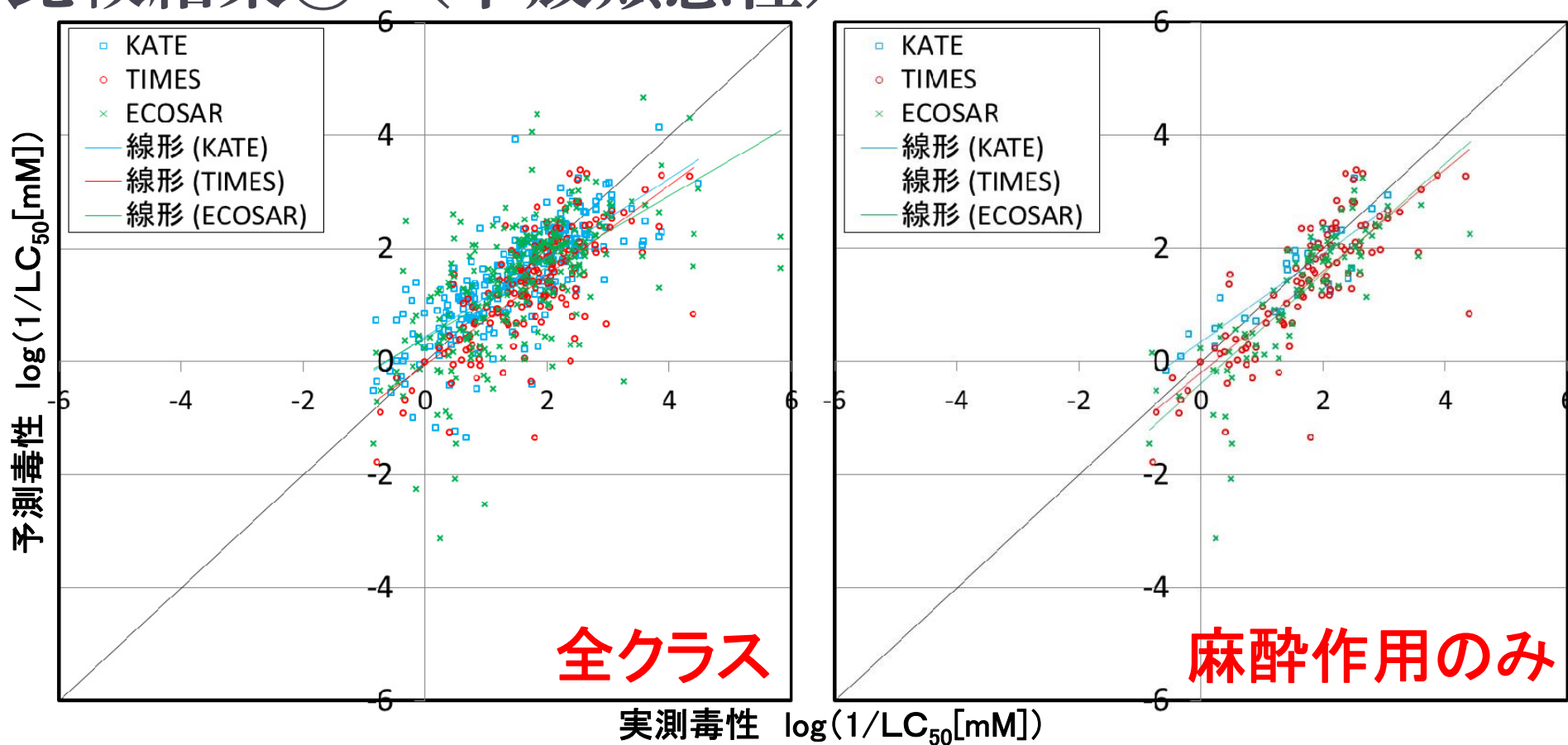
麻酔作用のみ: 麻酔作用で説明できる
クラスのみ予測

- KATE: Neutral Organics
- TIMES: basesurface narcotics, narcotic amines
- ECOSAR: Neutral Organics , Neutral Organics-acid

- RMSE: 平均二乗誤差
- R^2_{adj} : 自由度調整済決定係数

モデル		R^2_{adj}	RMSE
KATE	全クラス	0.78	0.60
	麻酔作用	0.91	0.41
TIMES	全クラス	0.83	0.49
	麻酔作用	0.85	0.50
ECOSAR	全クラス	0.64	1.1
	麻酔作用	0.77	0.79

比較結果② (甲殻類急性)



麻酔作用のみ: 麻酔作用で説明できる
クラスのみ予測

- KATE: Neutral Organics
- TIMES: basesurface narcotics, narcotic amines
- ECOSAR: Neutral Organics, Neutral Organics-acid

- RMSE: 平均二乗誤差
- R^2_{adj} : 自由度調整済決定係数

モデル		R^2_{adj}	RMSE
KATE	全クラス	0.60	0.62
	麻酔作用	0.76	0.46
TIMES	全クラス	0.35	0.53
	麻酔作用	0.65	0.79
ECOSAR	全クラス	0.36	1.4
	麻酔作用	0.68	0.46

KATE(KAshinhou Tool for Ecotoxicity)概要



予測の対象となる化学物質

- ▶ 一般工業化学物質
 - ▶ 特定の生物活性がある物質は予測不可(例: 農薬、医薬品)
- ▶ 単一の構造を持つ物質
- ▶ 分子量1000未満の物質
- ▶ 水環境中で安定な物質
- ▶ 水環境中で溶存態として溶解している物質
- ▶ 有機化学物質(有機金属は除く)
 - ▶ 無機化合物は予測不可
- ▶ 塩の場合は酸で代替
 - ▶ 例: カルボン酸ナトリウム⇒カルボン酸で予測
- ▶ タンパク質等との反応性を有しない物質

上記物質でも、ドメイン外となった予測結果は利用不可



KATE開発の経緯

2012年4月 改正化審法2段階目施行(スク評価・リスク評価)

2011年3月 KATE 2011公開 (PASは開発休止中)

2009年3月 インターネット版に加え
スタンドアロン(Windows)版を公開

2008年1月 試用版 (KATE Ver0.1) 公開

2007年7月 三省合同審議会に対し、
予測結果の提供を開始

2004年4月 環境省の請負業務として研究・開発
開始(2004年度～2012年度)

2003年 化審法改正(生態)

有害性クラス

1: $PNEC \leq 0.001$ mg/L
2: $0.001 < PNEC \leq 0.01$ mg/L
3: $0.01 < PNEC \leq 0.1$ mg/L
4: $0.1 < PNEC \leq 1$ mg/L
外: $PNEC > 1$ mg/L

第三種監視化学物質
 $NOEC \leq 0.1$ mg/L

生態毒性予測システムKATE

Knowledge-Driven QSAR Model

Quantitative Structure-Activity Relationship (QSAR)

定量的構造活性相関

化学物質の
構造上の特徴から事前に
クラスに分類する

- ★ Multiple QSAR models
- ★ battery QSAR models
- ★ combine classes
- ★ check the fragments in the training set of model

相関

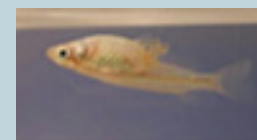
ULR

MLR

logP等

化学物質規制で要求
される毒性値

・ 魚類毒性値



・ ミジンコ毒性値

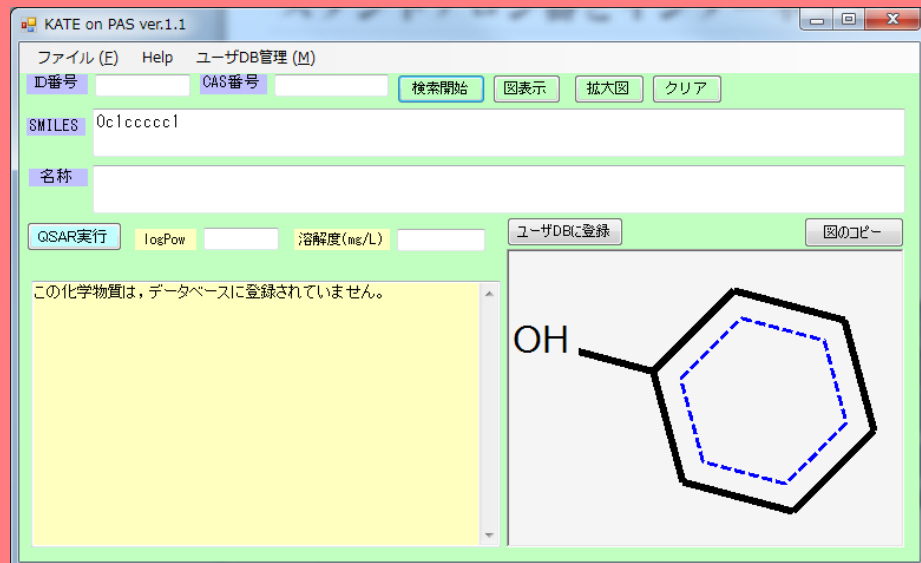


・ 藻類毒性値



(・ ユスリカ毒性値)

スタンドアロン版とインターネット版



KATE on PAS
(スタンドアロン版)

https://kate.nies.go.jp/nies/user_forecast_result_0 = 生態毒性予測システム

国立環境研究所 > 環境リスク研究センター > 生態毒性予測システム

Logout > Input molecule > View result > Verify QSAR Password Logout

生態毒性予測システム (KATE) ようこそ hasunuma4さん

CAS 未入力

物質名 未入力

SMILES c1ccccc1O

分子量 94.11

水溶解度 (mg/L) 入力(実測値) 更新

LogP DB(実測値) 1.47

生分解性 入力値 未入力

構造より予測 化学物質特性予測システム

HO

QSAR結果

クラス	エンドポイント	LogP	予測結果 (mg/L)	判定	溶解度	
				構造C ※2	LogP ※3	溶解度 ※4
amines aromatic or phenols4	Fish 96hr LC50	DB実測値(LogPstar)	44	○	○	未入力
amines aromatic or phenols4	Daphnid 48hr EC50	DB実測値(LogPstar)	10	○	○	未入力

KATE on NET
(インターネット版)

FITS (Fragment Identification by Tree Structure)

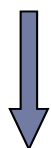
KATEシステム

曖昧でないアルゴリズム

QSAR予測

化学物質の構造 (SMILES)の入力

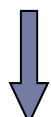
- SMILES: 化合物の分子構造を線形表記した識別子



- FITS: 部分構造の取得
Aromatic c-OH



- QSAR式を選択



amines aromatic or phenols4
 $\log(I/LC50) = 0.72\log P - 0.72$

- 記述子の取得

- ・実測値
- ・スタンダード版: U.S. EPA KOWWIN
- ・NET版: BioByte社 ClogP

部分構造に分解

モデル開発

構造分類

- ・部分構造の定義
- ・分類ルール (部分構造の組合せとその優先順位)



QSAR式を構成する
参照物質の部分構造
と比較



(部分構造)

○

適応可能領域 (記述子)

○

定義されたエンドポイント

Fish LC50 = 44 (mg/L)

当てはまりの良さ、
予測性の指標

計画(plan)

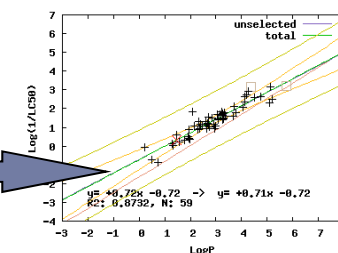
参照物質データ
・環境省生態毒性試験結果
・U.S. EPA アットヘッド・
ミノ データベース

実行(do)

評価(check)

改善(act)

検証データ
・環境省既存点検結果
・新規届出
既知見(反応性)

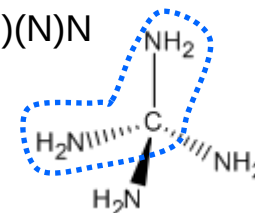


PASとは

- ▶ PAS (Platform for Assessment from Structure)^注は・・・
 - ▶ 構造分類に基づく物性や毒性を予測するための独自のシステム
 - ▶ 部分構造の取得プログラム (FITS; Fragment Identification by Tree Structure)、構造図の表示・入力プログラムなどからなる統合システム
- ▶ FITSは部分構造の規定に独自の記号を使用
 - ▶ 主体部分は、1次元構造を基本としたFITS記述です。
F/0|2|1|/C=CNC=C/I|nC=O,3V3,3B3,2Cy,3Cy,4Cy,2Rs4,/|

例: NC(N)(N)Nの構造でNCNの構造の数を、目的に応じて1－6個まで定義できます。

SMILES: NC(N)(N)N



注: PASの開発は、2000～2002年度(H12～14)環境省環境研究総合促進費「環境中の複合化学物質による次世代影響リスクの評価とリスク支援に関する研究」の一環として大分大学で実施。また、「環境データの解析と環境中生物影響評価に関する研究」として、2005～2008年度(H17～20)には(独)国立環境研究所と大分大学との委託・共同研究として実施。

Presence of Substructures

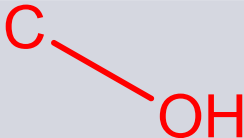
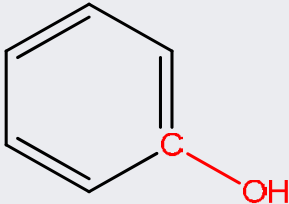
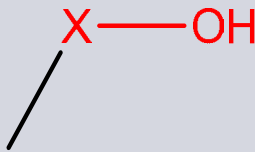
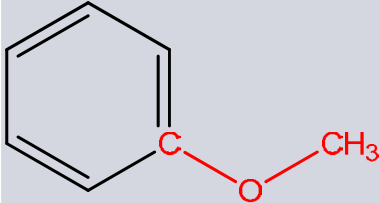
- ▶ 生物学的活性の可能性を示す部分構造と、その有無に基づいたQSARsが開発されている。
- ▶ 生態毒性では部分構造の評価が可能な大きなデータベースが入手可能であり、分類することにより、予測の誤差を減らす可能性がある。
- ▶ 一方、部分構造の取り扱いは困難。例えば、部分構造間の電子の相互作用を見込むことはできない。
- ▶ 予測する化学物質に新規な部分構造がある場合は、元のデータベースに存在していないため適切に評価できているか保証がない。



- ▶ 参照物質との比較によるドメイン判定が必要 (trade off)

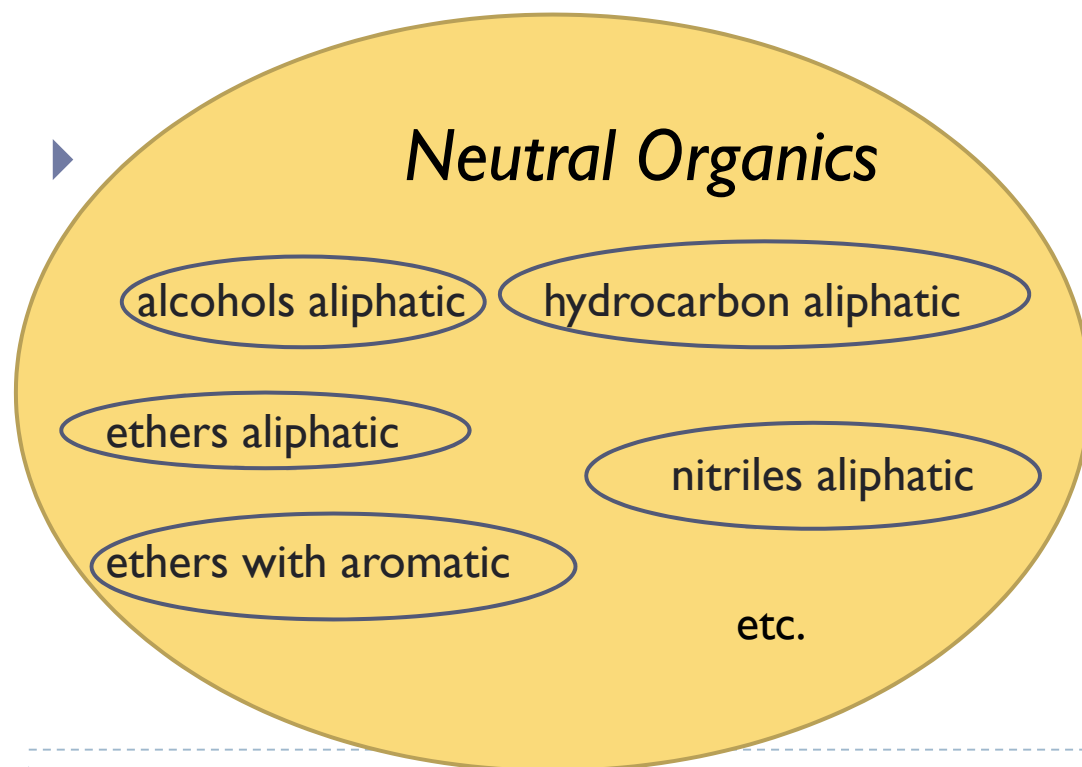


参照物質との構造比較に用いている部分構造の例

ID	説明	FITS	部分構造イメージ
5011	脂肪族C-OH	F/11/CO/2H1,/	
5012	芳香族c-OH	F/11/cO/2H1,/	
5013	X(炭素以外)-OH	F/11/?O/1?!;C;c;,2H1,/	
5014	脂肪族C-OMe	F/111/COC/3H3,/	
5015	芳香族c-OMe	F/111/cOC/3H3,/	
:	:	:	:

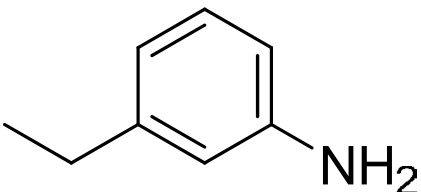
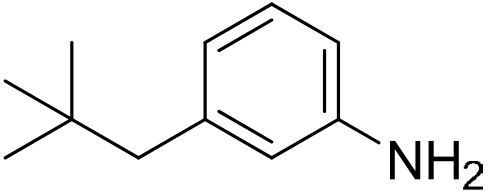
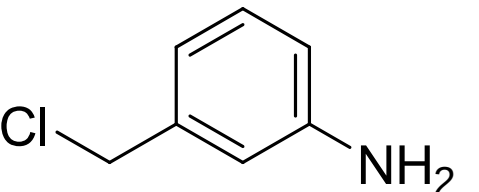
combine classes

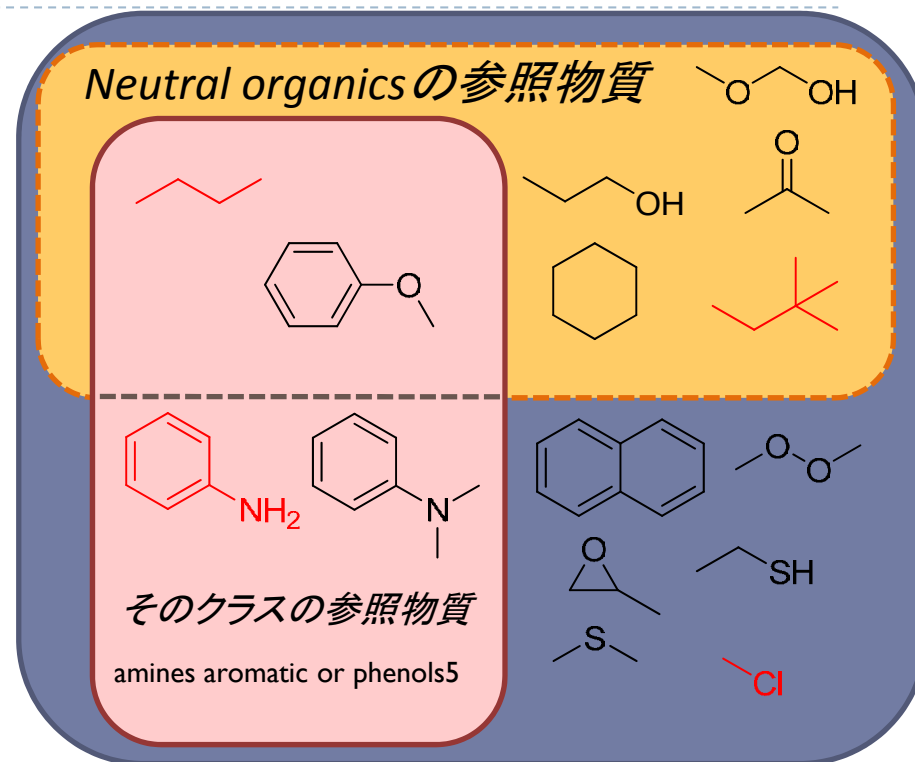
- ▶ Neutral Organicsクラス：脂肪族炭化水素、脂肪族・芳香族エーテル、脂肪族・芳香族ケトン、アルコールといった**単純な麻酔作用のみで毒性が説明できる**と考えられるクラスの物質を統合



Neutral Organics
⇒本クラスの参照物質の部分構造は反応性の毒性に寄与しないとし、他のクラスの構造C判定においても参照物質に含める。

部分構造によるクラス分類の適用範囲

	C(1)	C(2)
	YES 構造C判定：○	—
	NO 構造C判定：△	YES
	NO 構造C判定：×	NO



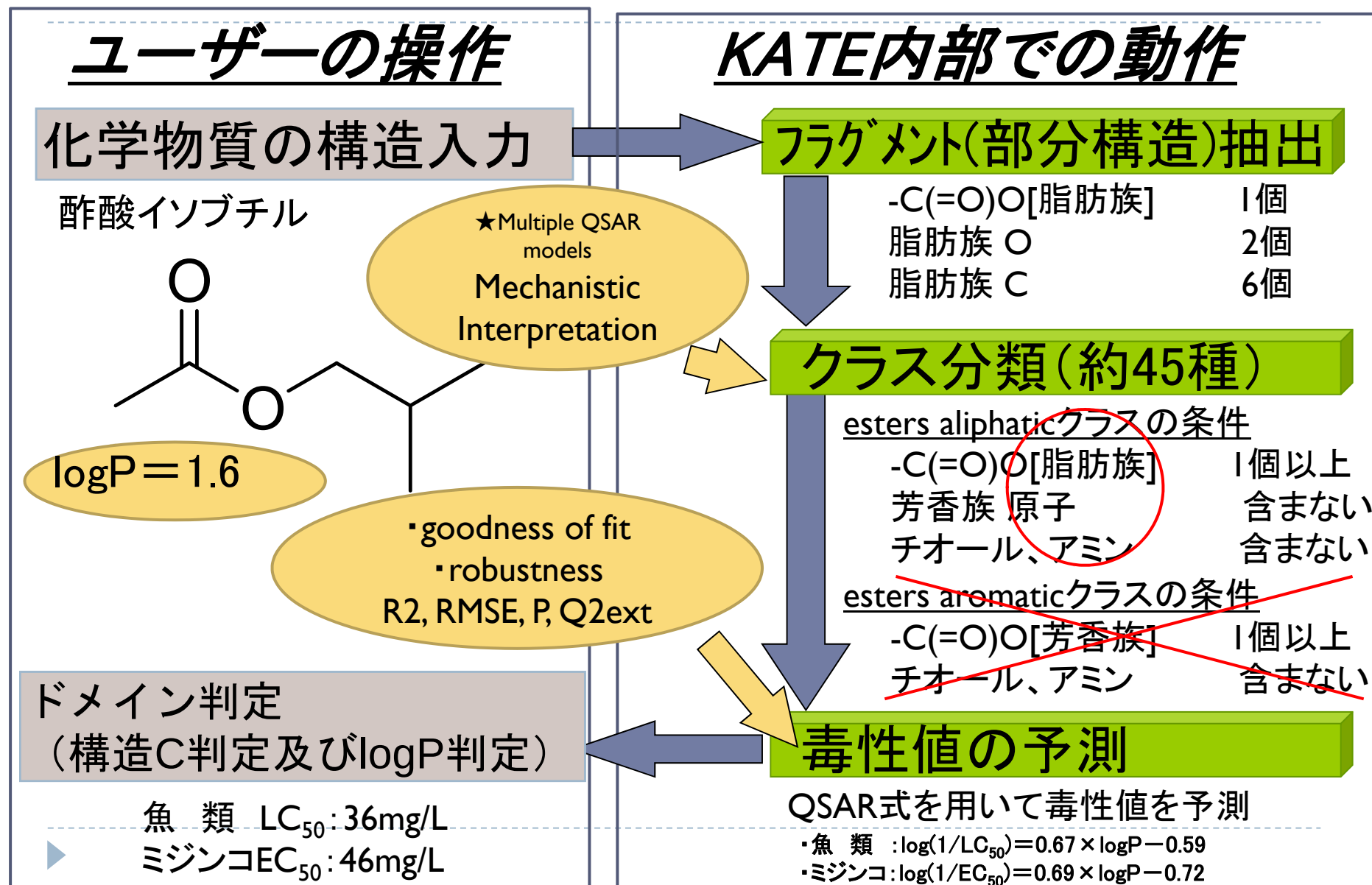
構造C 判定

参照物質の部分構造のリスト

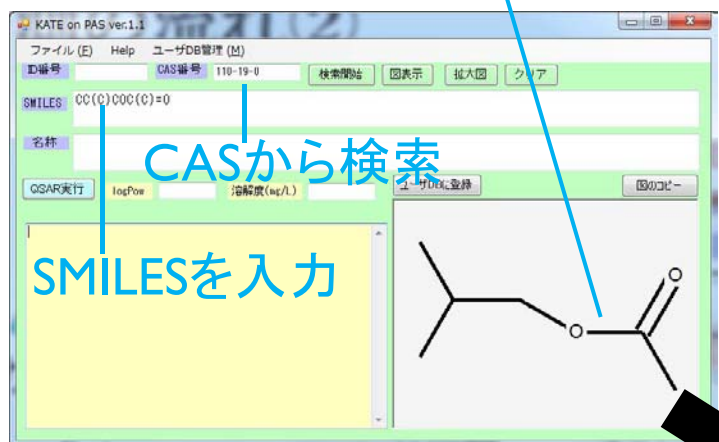
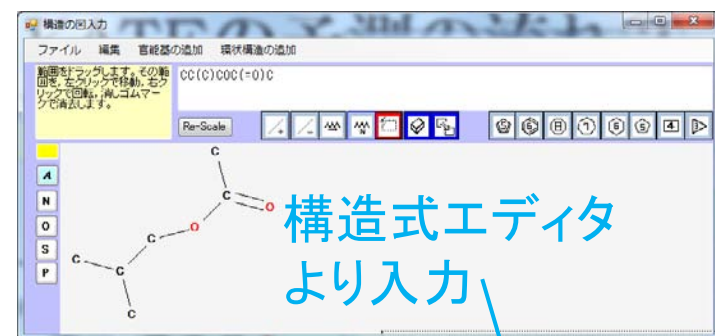
C(1): 予測した化学物質の**全ての部分構造**が、「そのクラス」の参照物質の部分構造リストに含まれるか。

C(2): 予測した化学物質の**全ての部分構造**が、「そのクラス」又はNeutral Organicsの参照物質の部分構造リストに含まれるか。

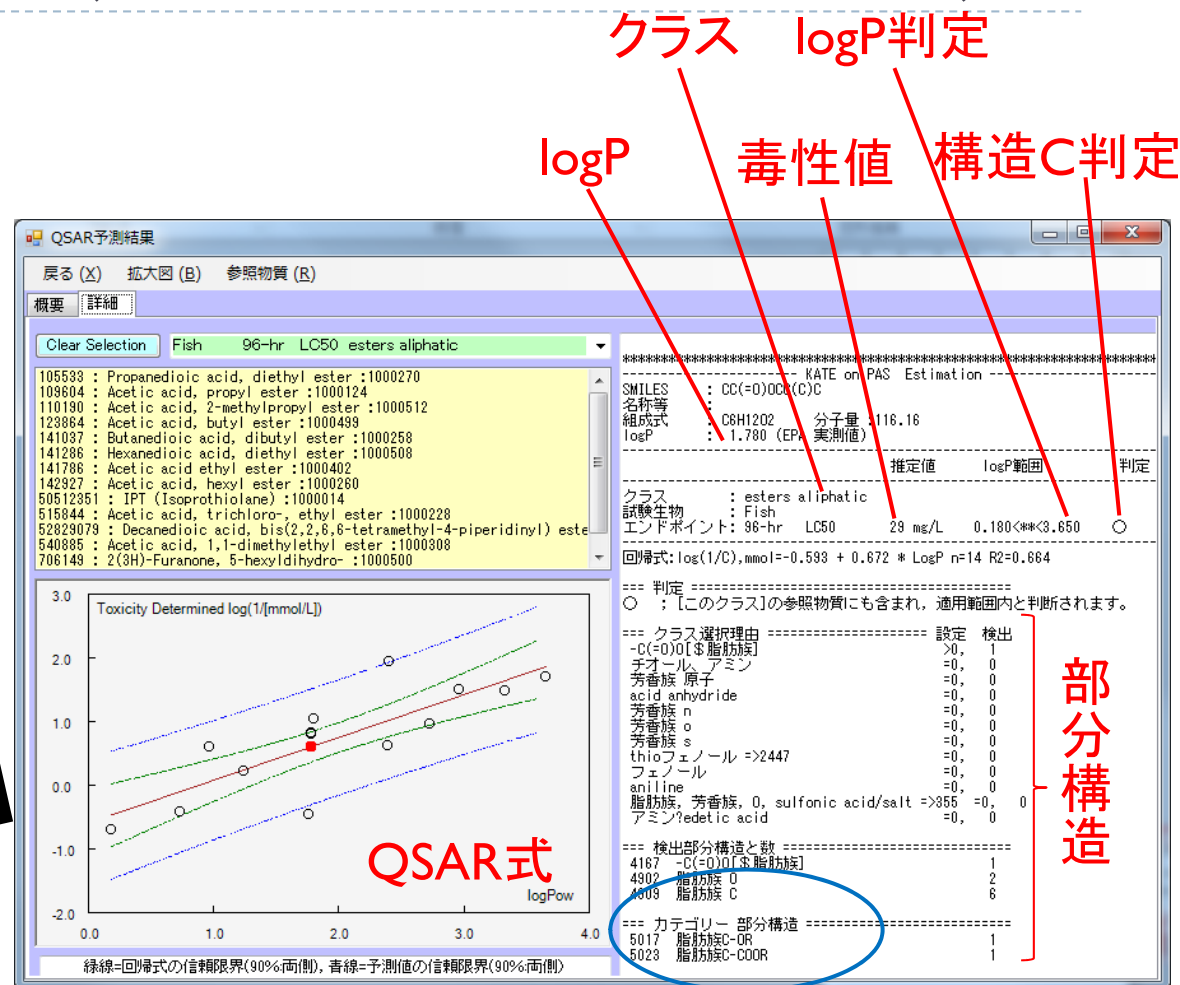
KATEの予測 (Battery QSAR models)



KATE on PAS(スタンドアロン版)



化学物質の構造入力



Battery QSAR models (Score)

予測
不可

構造
または
作用
機序

Name	Score
N or P cations	37
metals	36
Acrylate / acrylic acids	35
conjugated systems1	34
A-group	33
amines aromatic or phenols	32
halides reactive	31
dinitrobenzenes	30
epoxides	29
pyrethroids	28
carbamates	27
conjugated systems2	26
barbitals	25
hydrazines or polyamines	24
phenols	23
primary amines aliphatic or aromatic	22
amines aromatic	21
methacrylates	20
esters phosphate	19

Name	Score
thiols or disulfides	18
halides low-reactive	17
esters aliphatic	16
ethers with aromatic	15
amides or imides	14
aldehydes	13
secondary or tertiary amines	12
esters aromatic	11
ethers aliphatic	10
alcohols aliphatic	9
hydrocabons aliphatic	8
sulfides	7
nitrobenzenes	6
hydrocarbons aromatic	5
ketones	4
phosphates	3
nitriles aliphatic	2
acids	1
Unclassified	0

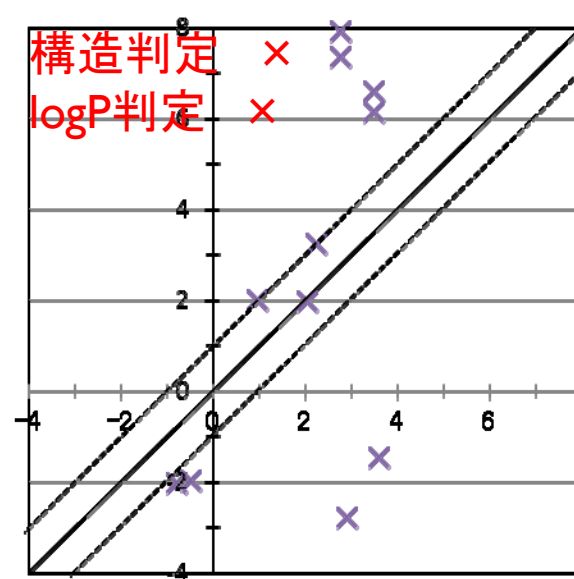
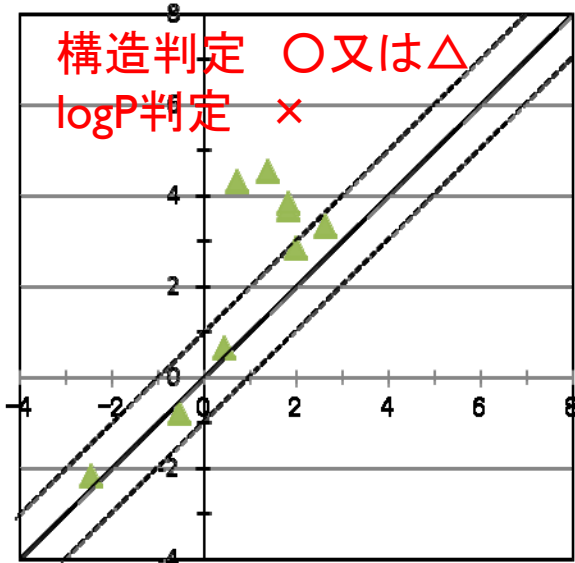
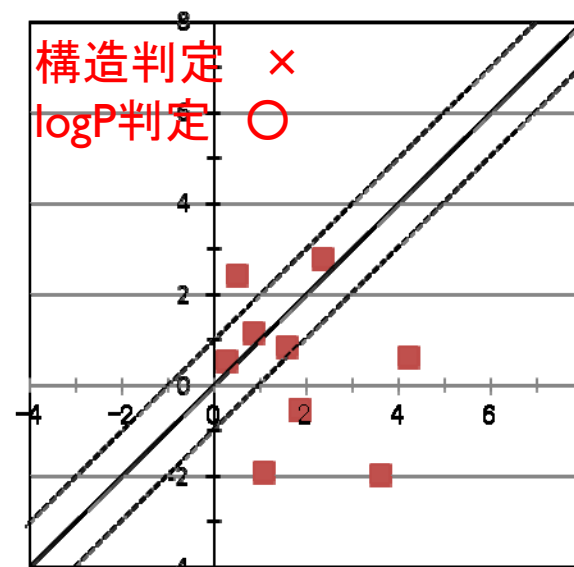
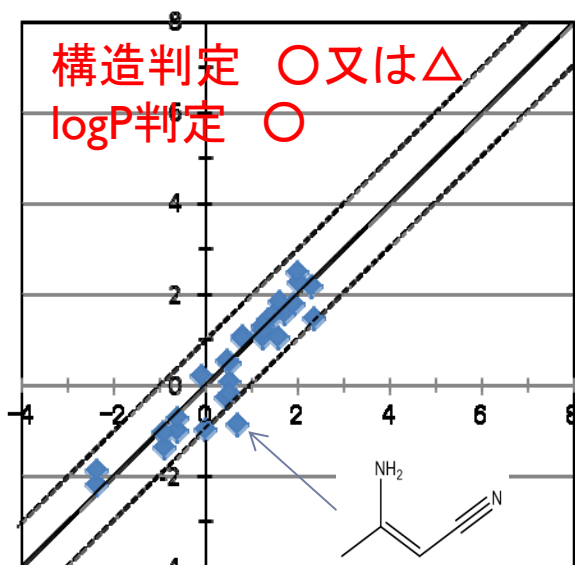
一般化学物質の有害性情報を活用した 外部バリデーション結果

- ▶ バリデーションに用いたQSAR式
 - ▶ KATE2011年公開版
 - ▶ 精度が高いと考えられるQSAR式のみ使用 ($r^2 > 0.7$ 、 $p(\text{slope}) < 0.05$)
- ▶ バリデーションに用いた毒性情報
 - ▶ KATE2011公開後に報告された環境省が実施した生態毒性試験結果（魚類急性毒性試験、ミジンコ遊泳阻害試験）
<http://www.env.go.jp/chemi/sesaku/02e.pdf>
 - ▶ H25.07.19 化学物質審査小委員会（第135回）の「資料4-2 生態影響に関する優先度判定案」における魚類・ミジンコ類急性毒性値のうちKATE2011の参照物質として使用されていない物質
http://www.env.go.jp/council/05hoken/y051-135-1/mat04_2.pdf



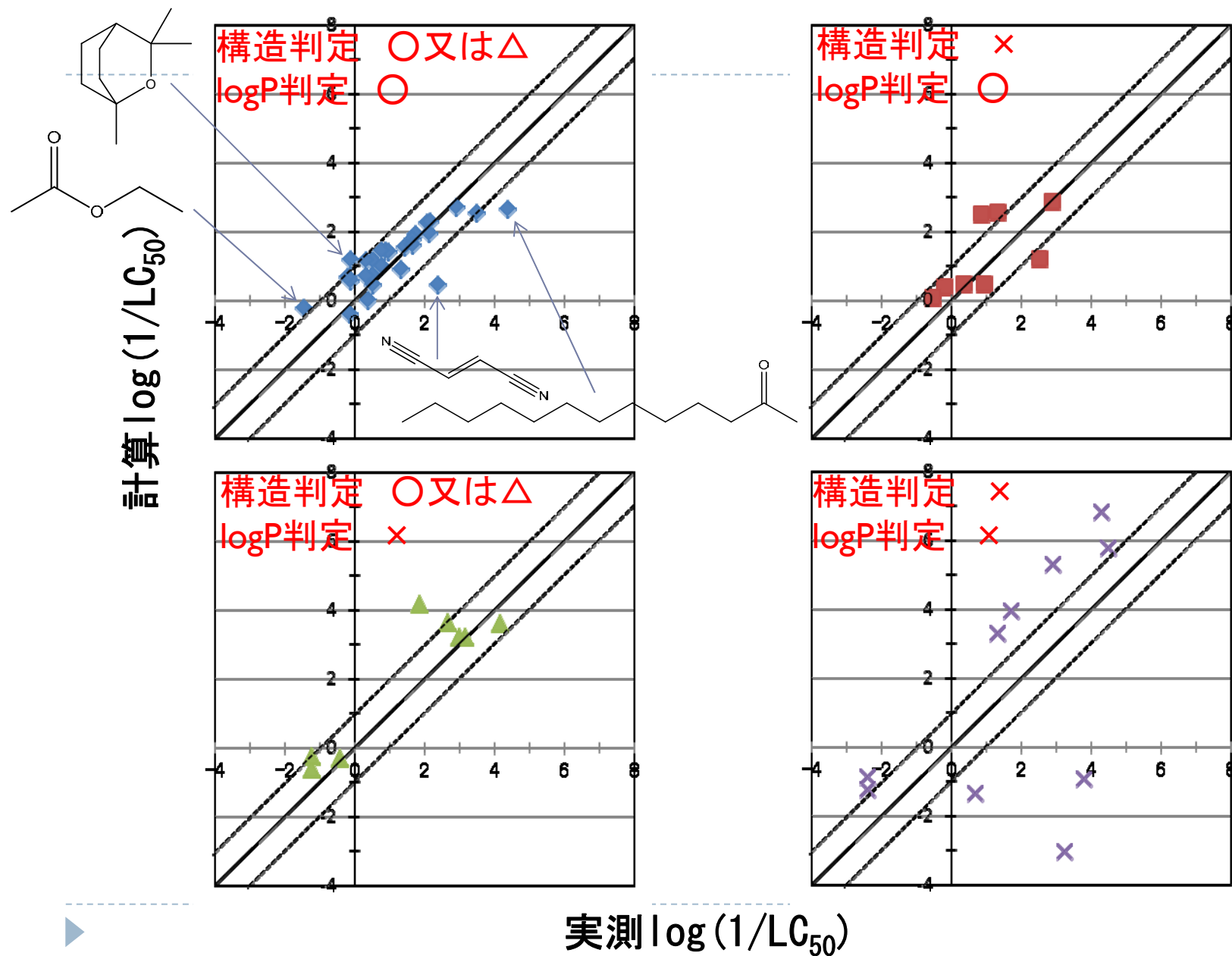
魚類バリデーション結果(n=62)

計算 $\log(1/LC_{50})$



実測 $\log(1/LC_{50})$

甲殻類バリデーション結果(n=53)



KATEの活用検討状況



化審法スクリーニング評価におけるQSAR、 カテゴリーアプローチの活用検討

- ▶ QSAR やカテゴリーアプローチの導入については、スクリーニング評価作業の中のどのような場面で活用可能か等を早急に検討し、活用可能と考えられる部分については、一般化学物質のスクリーニング評価の実施に合わせて試行することを目指すとしたが、より具体的な検討ステップは以下のとおり考える。
- ▶ 1. スクリーニング評価において、どのような場面でQSAR やカテゴリーアプローチが活用可能かを検討する。その際、①有害性を過小に評価しないこと。②効率的で低負荷なスクリーニング評価の実施に貢献すること等を念頭に検討を行う。
- ▶ 2. 人健康、生態に対する候補QSAR モデルについて、新規化学物質、既存点検等用いた試験データとの検証を進めてきた推計成績(正解率、統計データ)をまとめる。
- ▶ 3. 1. で活用すべきと判断した場面において、2. の推計成績を加味し、利用可能なQSAR モデルやカテゴリーを利用するものとする。具体的に想定される活用事例としては、「有害性情報が得られない場合の代用」、「評価を行う順序付け」等が挙げられる。なお、後者の活用例等は、推計成績が必ずしも高い必要がないと考えられることから、積極的なQSAR やカテゴリーアプローチの活用を行う。
- ▶ 4. このような実績を積み上げながら、国際動向や国内外のQSAR モデルの開発動向やカテゴリーアプローチの活用動向も注視し、一層、適用範囲を広げていく。
- ▶ なお、1～4の検討については別途3省で毒性等の専門家の意見を踏まえつつ検討を進め、一般化学物質のスクリーニングの開始までに3省の審議会では具体的な利用方法を取りまとめるものとする。

スクリーニング評価の基本的な考え方における(三省合同審議会、平成22年10月8日)より



化審法スクリーニング評価における検討事例

生態影響に係る優先度「中」区分からの優先評価化学物質選定について(抜粋)

物質名	有害性 クラス	PNEC (mg/L) (A)/(B)	最小値 (mg/L) (A)	Ufs (B)	藻類(mg/L)				ミジンコ類(mg/L)				魚類(mg/L)			
					急性 毒性値 (EC50)	EC50/AC R	慢性 毒性値 (NOEC)	NOEC/UF (種間外 挿)	急性 毒性値 (EC50)	EC50/AC R	慢性 毒性値 (NOEC)	NOEC/U F(種間外 挿)	急性毒 性値 (LC50)	LC50/A CR	慢性 毒性 値 (NOE C)	NOEC/U F(種間外 挿)
オクタデシルアミン	1	0.000013	0.13	10000	0.12	0.006	0.029	0.0029	0.13	0.0013						
N, N-ジメチルドデ カン-1-イルアミン	1	< 0.000052	< 0.0026	50	0.014	0.0007	< 0.0026	< 0.00052	0.083	0.00083	0.036	0.0072	0.57	0.0057		
ココアルキルアミン	1	0.000045	0.045	1000	0.17	0.0085	0.071	0.0071	0.045	0.00045			0.16	0.0016		

三省合同審議会(平成25年7月19日)より

魚類急性毒性:ミジンコよりも毒性弱い可能性?

魚類急性毒性ありと見なす場合(エキスパートジャッジ)

- ・UF: 10000(急性二種)⇒1000(急性三種)
- ・PNEC: 0.000013 mg/L⇒ 0.00013 mg/L
- ・スクリーニング評価結果: 優先評価化学物質相当⇒一般化学物質相当

オクタデシルアミンKATE予測結果:

- ・ミジンコ48時間EC₅₀: 0.00091 mg/L
- ・魚類96時間LC₅₀: 0.037 mg/L
- ・ドメイン判定(構造): 内
- ・ドメイン判定(logP): 外
(本物質logP: 7.7, 参照物質log上限: 5.4)

今回はドメイン外であり活用不可
次回以降のスクリーニング評価で
活用可能か、引き続き検討を実施

化学物質の環境リスク初期評価における活用：検討中

- ▶ QSARの生態リスク初期評価への活用方法について、活用の限界等について論点を整理した上でルール化について検討中。

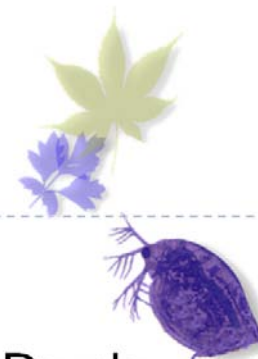
表 水生生物に対する毒性値の概要

生物群	急性	慢性	毒性値 [μg/L]	生物名	生物分類／和名	エンドポイント ／影響内容	ばく露 期間[日]	試験の 信頼性	採用の 可能性	文献No.
藻類	○		640	Pseudokirchneriella subcapitata	緑藻類	EC ₅₀ GRO(RATE)	2	B	B	I)-106416
	○		7,100	Skeletonema costatum	珪藻類	EC ₅₀ GRO	4	D	C	I)-9607
	○		52,900	Pseudokirchneriella subcapitata	緑藻類	EC ₅₀ GRO	4	D	C	I)-9607
甲殻類	○		1,480	Americamysis bahia	アミ科	LC ₅₀ MOR	4	D	C	I)-9607
	○		>530,000	Daphnia magna	オオミジンコ	LC ₅₀ MOR	2	C	C	I)-5184
魚類			69	Jordanella floridae	キプリノドン科 (仔魚)	NOEC GRO	28	B	C	I)-140
	○		>89	Pimephales promelas	ファットヘッドミノ	LC ₅₀ MOR	4 (止水式)	C	C	I)-17138
		○	90	Cyprinodon variegatus	キプリノドン科 (胚)	NOEC MOR (仔魚)	～ふ化後 28	B	B	I)-9953
	○		305	Pimephales promelas	ファットヘッドミノ	LC ₅₀ MOR	4 (流水式)	D	C	4)-2012253
:	:	:	:	:	:	:	:	:	:	:



化学物質の環境リスク評価 第11巻 1,2,4,5-テトラクロロベンゼンより

ご静聴ありがとうございました



参考文献

- 1 A. Furuham, T. Toida, N. Nishikawa, Y. Aoki, Y. Yoshioka, H. Shiraishi, Development of an ecotoxicity QSAR model for the KAshinhou Tool for Ecotoxicity (KATE) system, March 2009 version, SAR QSAR Environ. Res., 21 (2010), pp. 403-413.
- 2 A. Furuham, K. Hasunuma, Y. Aoki, Y. Yoshioka and H. Shiraishi, Application of chemical reaction mechanistic domains to an ecotoxicity QSAR model, KAshinhou Tool for Ecotoxicity (KATE), SAR QSAR Environ. Res., 22 (2011), 505-523.
- 3 A. Furuham, Y. Aoki, and H. Shiraishi, Development of ecotoxicity QSAR models based on partial charge descriptors for acrylate and related compounds, SAR QSAR Environ. Res., in press.
- 4 吉岡義正「QSAR用プラットフォーム(PAS)作成のノウハウ-1 SMILES式からの要素の抽出-」環境毒性学会誌 Vol. 11 (2008), pp. 33-40 .
- 5 吉岡義正「QSAR用プラットフォーム(PAS)作成のノウハウ-2 構造図の描画-」環境毒性学会誌 Vol. 12 (2009) , pp. 107-112.
- 6 吉岡義正「QSAR用プラットフォーム(PAS)作成のノウハウ-3 部分構造の取得法-」環境毒性学会誌 Vol. 12 (2009), pp. 113-122.
- 7 古濱彩子、白石寛明「化学物質の生態毒性予測システムKATEとQSAR」日本化学会情報化学部会誌 Vol. 30 (2012), pp.42-45.

▶ KATEの詳細・ダウンロード先：<https://kate.nies.go.jp>