

生態毒性QSARモデルの解説 (KATE、OECD QSAR Toolbox デモンストレーション)

大分大学教育福祉科学部

吉岡 義正(大阪会場のみ)

(独)国立環境研究所 環境リスク研究センター

蓮沼 和夫

2010年1月25日(月)東京会場

2010年1月28日(木)大阪会場

はじめに

QSARで得られた予測結果は、化審法の届出に必要な生態毒性試験結果として利用することは出来ません

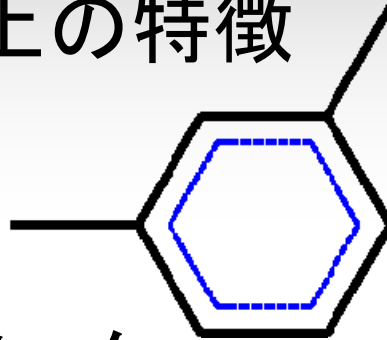
QSARとは

■ Quantitative Structure-Activity Relationshipの略

■ 化学物質の構造上の特徴

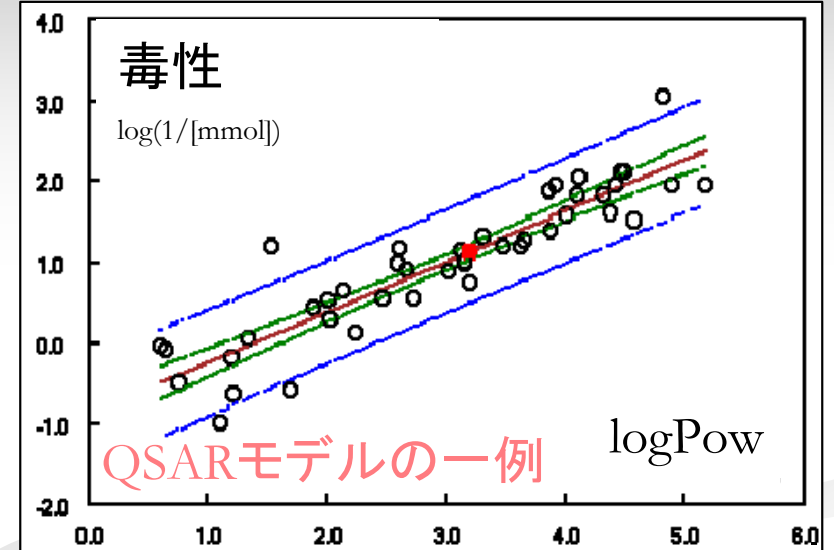
■ 脂肪族C: 2個

■ 芳香族原子: 6個



■ 物理化学的パラメータ

■ LogP_{ow} (水-オクタノール分配係数): 3.2



相関関係

■ 生物学的活性 (毒性等)

■ 構造活性相関 (SAR)

■ 好氣的分解性: 難分解性

■ 定量的構造活性相関 (QSAR)

■ 魚類 LC_{50} : 7.8mg/L

KATEの解説

生態毒性予測システムKATE

- KATE (KAshinhou Tool for Ecotoxicity) は・・・
 - 化学物質の部分構造から毒性値を予測：
 - 魚類急性毒性試験における半数致死濃度(LC₅₀)
 - ミジンコ遊泳阻害試験における半数影響濃度(EC₅₀)

2009年3月

- スタンドアロン版「KATE on PAS」公開
- Web版「KATE on NET」公開

2008年1月

試用版(KATE Ver0.1)公開

2007年7月

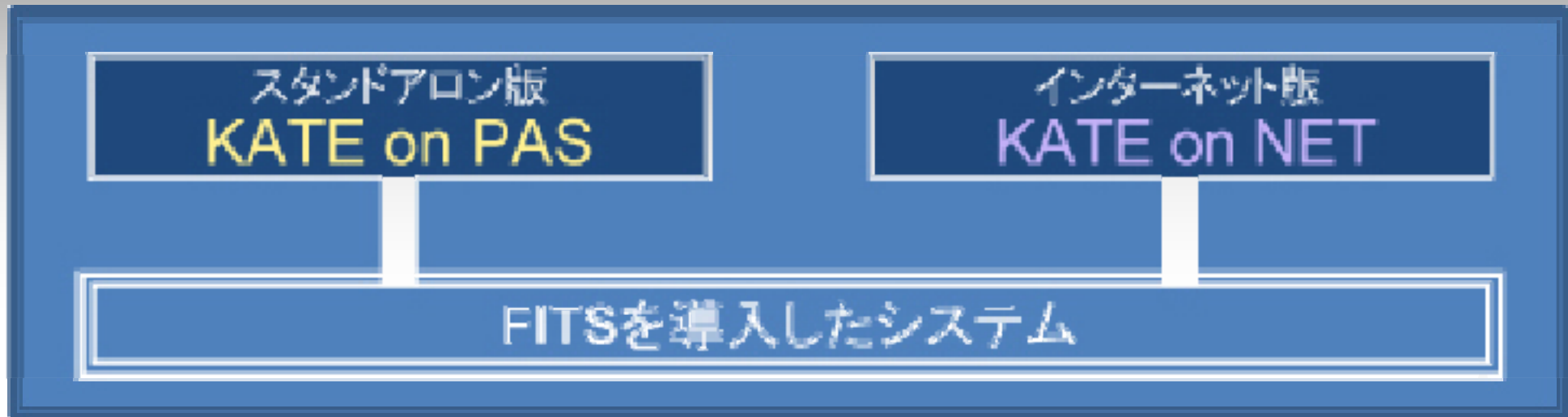
3省合同審議会※に対し、
予測結果の提供を開始

2004年

環境省の請負業務として研究・開発
開始(2004年度～2009年度)

注)※: 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

2つのKATE



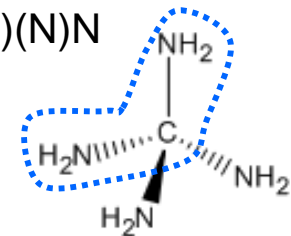
- 秘密保持の問題、透明性の確保、ライセンス上の問題からスタンドアロン版を開発
 - 両者とも予測結果は同一
 - ただし、 $\log P_{ow}$ 予測アルゴリズムが両者で異なる
 - ➔ ユーザーが $\log P$ を入力しない場合、予測値に差異が生じる可能性がある
- スタンドアロン版: EPA KOWWIN インターネット版: ClogP

PAS

- PAS (Platform for Assessment from Structure)[※]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ルールを使用
 - 主体部分は、1次元構造を基本としたFITS記述です。
F/01211/**C=CNC=C**/1JnC=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。

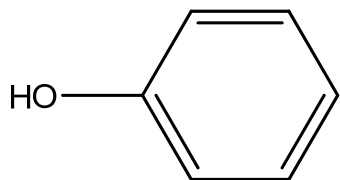
KATEでの予測方法

ユーザーの操作

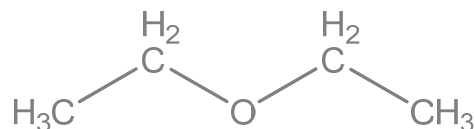
化学物質の構造
(SMILES)の入力

- SMILES: 化合物の分子構造を線形表記した識別子

- フェノール: c1ccccc1O



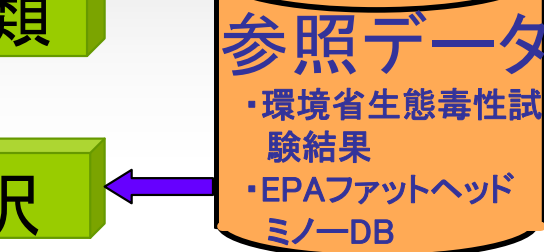
- ジエチルエーテル:
CCOCC



KATE内部での動作

クラスの分類

QSAR 式の選択



← logPow代入

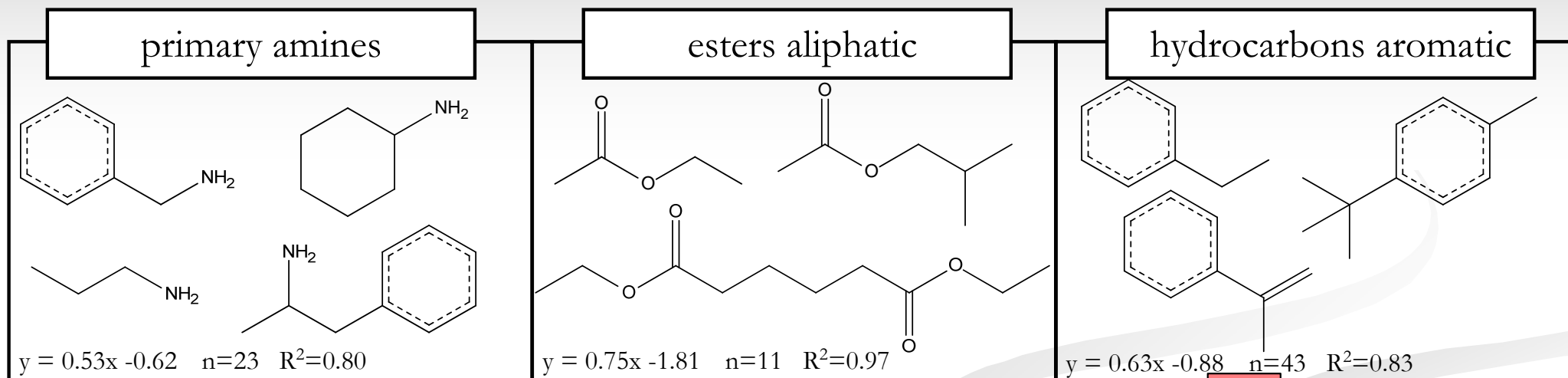
予測毒性値を出力

- 化学物質の構造に基づくクラス分類した線形式を使用
- 水一オクタノール分配係数 (logPow) の値から単相関で毒性値を予測

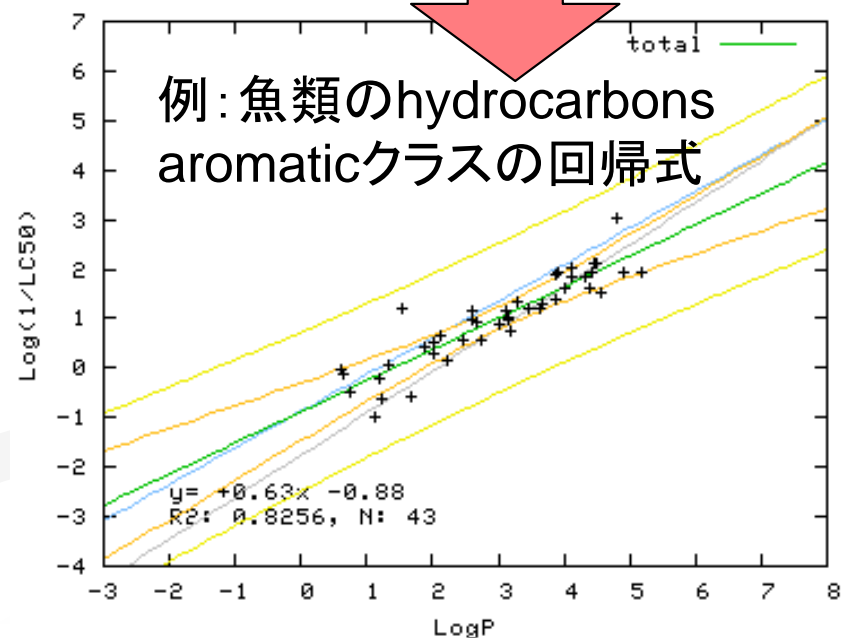
KATEのクラスについて

- 魚類・甲殻類合わせ約80種類のQSAR式・クラスが存在

【【魚類クラスの参照物質一例】】



	魚類	甲殻類
式の数(参照データ 2以上)	41	36
データ数が十分な式の数 (参照データ 5以上)	34	20
信頼できる式の数 (参照データ 5以上 かつR ² が0.7以上)	19	11



Neutral Organicsクラスについて

- 脂肪族炭化水素、スルホキシド、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分子種の一覧が生物種ごとに用意されており、これらはNeutral Organicsというクラスとして再定義されています。
- KATEでNeutral Organicsに分類されるクラス群：
 - alcohols or ethers aliphatic, ethers aliphatic , ethers aromatic , hydrocabons aliphatic , ketones , nitriles aliphatic , phosphate , sulfoxides

スタンドアロン版KATE on PAS 入力画面

KATE on PAS

ファイル (F) Help ユーザDB管理 (M)

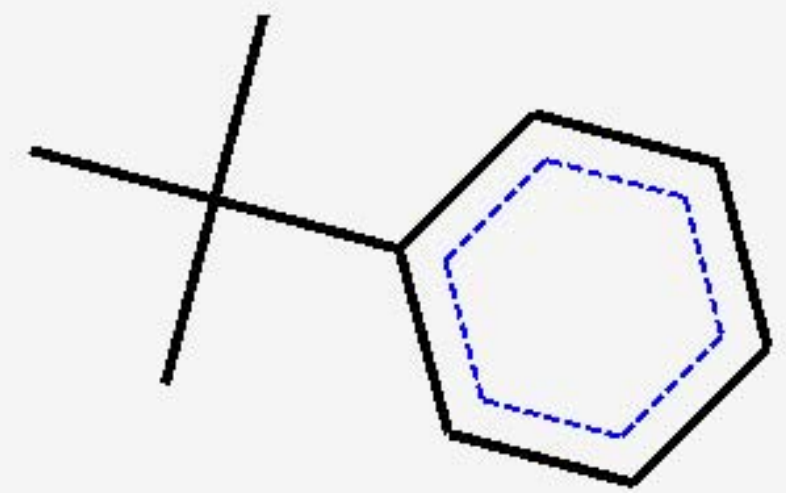
ID番号 CAS番号

SMILES ←SMILES

名称

logPow 溶解度 (mg/L)

logPow実測値(EPA)=4.11 (使用)
logPow推定値(EPA)=3.90



The image shows a screenshot of the KATE on PAS software interface. The main window has a blue title bar and a menu bar with 'ファイル (F)', 'Help', and 'ユーザDB管理 (M)'. Below the menu bar, there are input fields for 'ID番号' and 'CAS番号', followed by buttons for '検索開始', '図表示', '拡大図', and 'クリア'. The 'SMILES' field contains the string 'CC(C)(C)c1ccccc1' with a red arrow pointing to it and the text '←SMILES'. Below this is an empty '名称' field. Further down, there are buttons for 'QSAR実行', 'logPow', and '溶解度 (mg/L)', each followed by an input field. To the right of these are buttons for 'ユーザDBに登録' and '図のコピー'. At the bottom left, a yellow box displays 'logPow実測値(EPA)=4.11 (使用)' and 'logPow推定値(EPA)=3.90'. On the right side, a chemical structure of 1,1,1-trimethylbenzene is shown with a dashed blue outline.

KATE on PAS出力(概要)画面

QSAR予測結果

戻る (X) 拡大図 (B) 参照物質 (R)

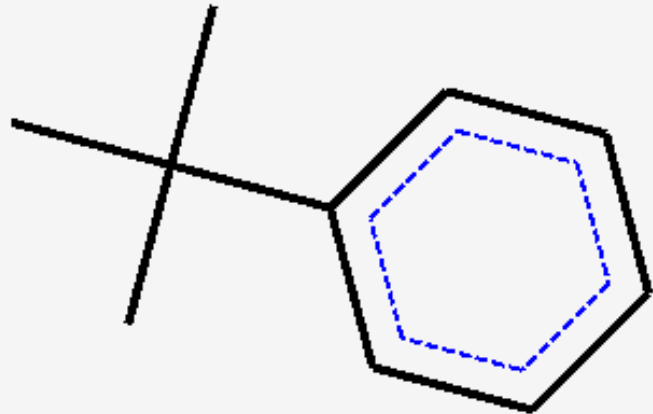
概要 詳細

分子量 134.22 使用logPow 4.11

溶解度

logPow実測値(EPA)=4.11 (使用)
logPow推定値(EPA)=3.90

KATEでは予め定義された部分構造の有無で、クラスの適用範囲を判断します。予測する化学物質のもつ部分構造のすべてが、[そのクラス]や[Neutral Organicクラス]の参照物質に含まれるかで、○、△、×に評価します。○、△の場合は[そのクラス]の適用範囲内と判断し



```

*****
----- KATE on PAS Estimation -----
SMILES      : CC(C)(C)c1ccccc1
名称等      :
組成式      : C10H14   分子量 :134.22
logP        : 4.110 (EPA 実測値)
-----
----- KATE 生態毒性推定値 -----
推定値      logP範囲      判定
-----
クラス      : hydrocarbons aromatic
試験生物    : Fish
エンドポイント: 96-hr LC50    2.6 mg/L    0.600<***<5.170    ○
-----
クラス      : hydrocarbons aromatic
試験生物    : Daphnid
エンドポイント: 48-hr EC50    1.1 mg/L    0.650<***<5.170    ○
-----
=== 判定 ===
○ ; [このクラス]の参照物質にも含まれ、適用範囲内と判断されます。

=== 検出部分構造と数 ===
4909 脂肪族 C          4
4910 芳香族 原子      6

=== カテゴリー 部分構造 ===
5056 脂肪族C-CH3      3
5063 芳香族t-Bu       1
5201 benzene          1
5236 Aromatics        1
    
```

予測毒性値

判定

KATE on PAS出力(詳細)画面

参照物質(QSAR式作成の根拠となった
実測毒性値が入手可能な物質)一覧

QSAR予測結果

戻る (X) 拡大図 (B) 参照物質 (R)

概要 詳細

Clear Selection Fish 96-hr LC50 hydro

100414 : Benzene, ethyl- :1000628
 100696 : Pyridine, 2-ethyl- :1000135
 100710 : Pyridine, 2-ethyl- :1000235
 104518 : Benzene, butyl- :1000341
 106423 : Benzene, 1,4-dimethyl- :1000612
 108383 : Benzene, 1,3-dimethyl- :1000490
 108883 : Benzene, methyl- :1000385
 108894 : Pyridine, 4-methyl- :1000239
 108996 : Pyridine, 3-methyl- :1000552
 109068 : Pyridine, 2-methyl- :1000020
 109977 : 1H-Pyrrole :1000417
 110009 : Furan :1000120
 110021 : Thiophene :1000916

KATE on PAS Estimation

SMILES : CC(C)(C)ClCCCC1
 名称等 :
 組成式 : C10H14 分子量 :134.22
 logP : 4.110 (EPA 実測値)

魚類QSAR式

	推定値	logP範囲	判定
クラス : hydrocarbons aromatic			
試験生物 : Fish			
エンドポイント: 96-hr LC50	2.6 mg/L	0.600<***<5.170	○

回帰式: $\log(1/C), \text{mmol} = -0.883 + 0.630 * \text{LogP}$ n=43 R2=0.826

=== 判定 ===
 ○ ; [このクラス]の参照物質にも含まれ、適用範囲内と判断されます。

=== クラス選択理由 ===

設定	検出
芳香族 原子	>0, 6
脂肪族 N NO	=0, 0
脂肪族 O	=0, 0
脂肪族 S	=0, 0
脂肪族 P	=0, 0
脂肪族 carbon-halogen	=0, 0
脂肪族 NO2	=0, 0
芳香族 n	=0, 0
芳香族 o	=0, 0
芳香族 s	=0, 0
芳香族 carbon-halogen	=0, 0

=== 検出部分構造と数 ===

4909 脂肪族 C	4
4910 芳香族 原子	6

=== カテゴリー 部分構造 ===

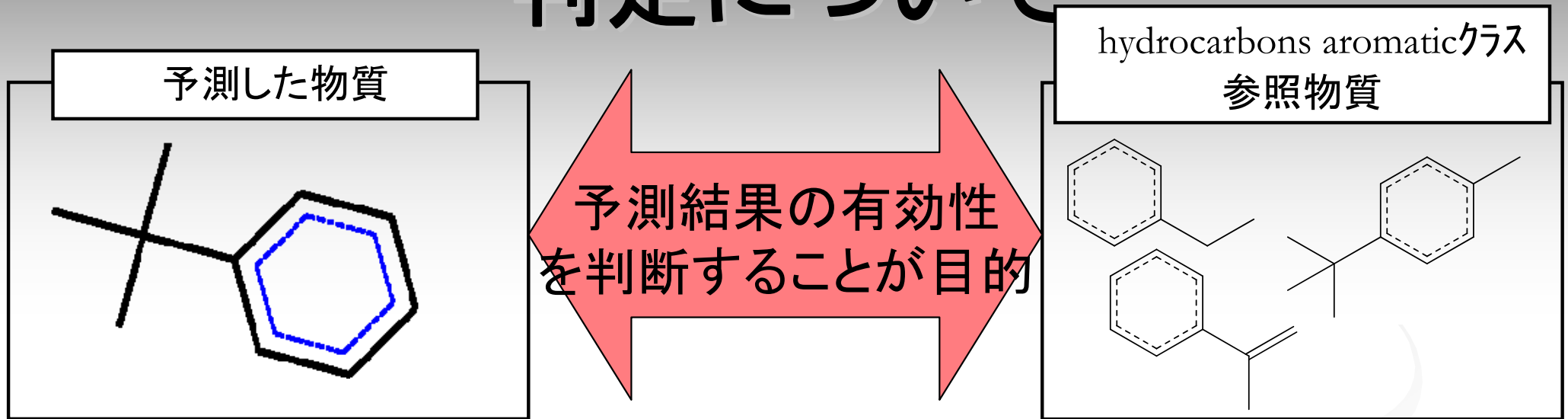
5056 脂肪族C-CH3	3
5063 芳香族c-t-Bu	1
5201 benzene	1
5236 Aromatics	1

Toxicity Determined $\log(1/[\text{mmol/L}])$

logPow

緑線=回帰式の信頼限界(90%:両側), 青線=予測値の信頼限界(90%:両側)

判定について



- 構造C判定: 予測する化学物質のもつ部分構造すべてが、
 - ○: [そのクラス]の参照物質にも含まれる。
 - △: [そのクラス]または[Neutral Organicクラス]の参照物質にも含まれる。
 - ×: [そのクラス]や[Neutral Organicクラス]の参照物質には含まれない部分構造がある。

として評価される。

- LogP判定: 予測した物質のlogPowがQSAR回帰式の有効範囲内に入っているか評価される。
 - KATE on PAS: 有効範囲外の場合、『>P』又は『<P』と評価
 - KATE on NET: 有効範囲内の場合『○』、範囲外は『×』と評価

KATE今後の課題

- 環境省生態影響試験の反映
 - 2008年3月版→2009年3月版
- クラス分類の修正
- logPow予測アルゴリズムの統一化 (EPA KOWWINに統一)
- logPowが大きい物質の毒性が過大評価される問題を修正

OECD (Q)SAR Application Toolbox の解説

OECDにおける取組

- 2002年3月 QSARの規制利用に関するICCAワークショップ(於ポルトガル・Setubal)
- 2004年11月第37回合同会合
 - ①QSARの規制利用に係るバリデーション原則に合意(Setubal原則の確認)
 - ②専門家グループ→QSARアドホックグループ(Ad Hoc Group on (Q)SARs)への改組提案(※QSAR専門家の集まりから、QSARを規制等に利用する者も参加する枠組みへ拡大)
- 2006年6月 第1回QSARアドホックグループ開催
- 2007年4月 第2回QSARアドホックグループ開催
- 2008年3月 (Q)SAR Application Toolbox1.0公開
- 2008年12月 (Q)SAR Application Toolbox1.1公開
- 2009年2月 第1回(Q)SAR Application Toolbox管理会合開催
- 2009年10月 第2回(Q)SAR Application Toolbox管理会合開催
- 2010年1月 (Q)SAR Application Toolbox1.1.02公開
- 2010年10月 (Q)SAR Application Toolbox2.0公開予定
- 2012年 (Q)SAR Application Toolbox3.0公開予定

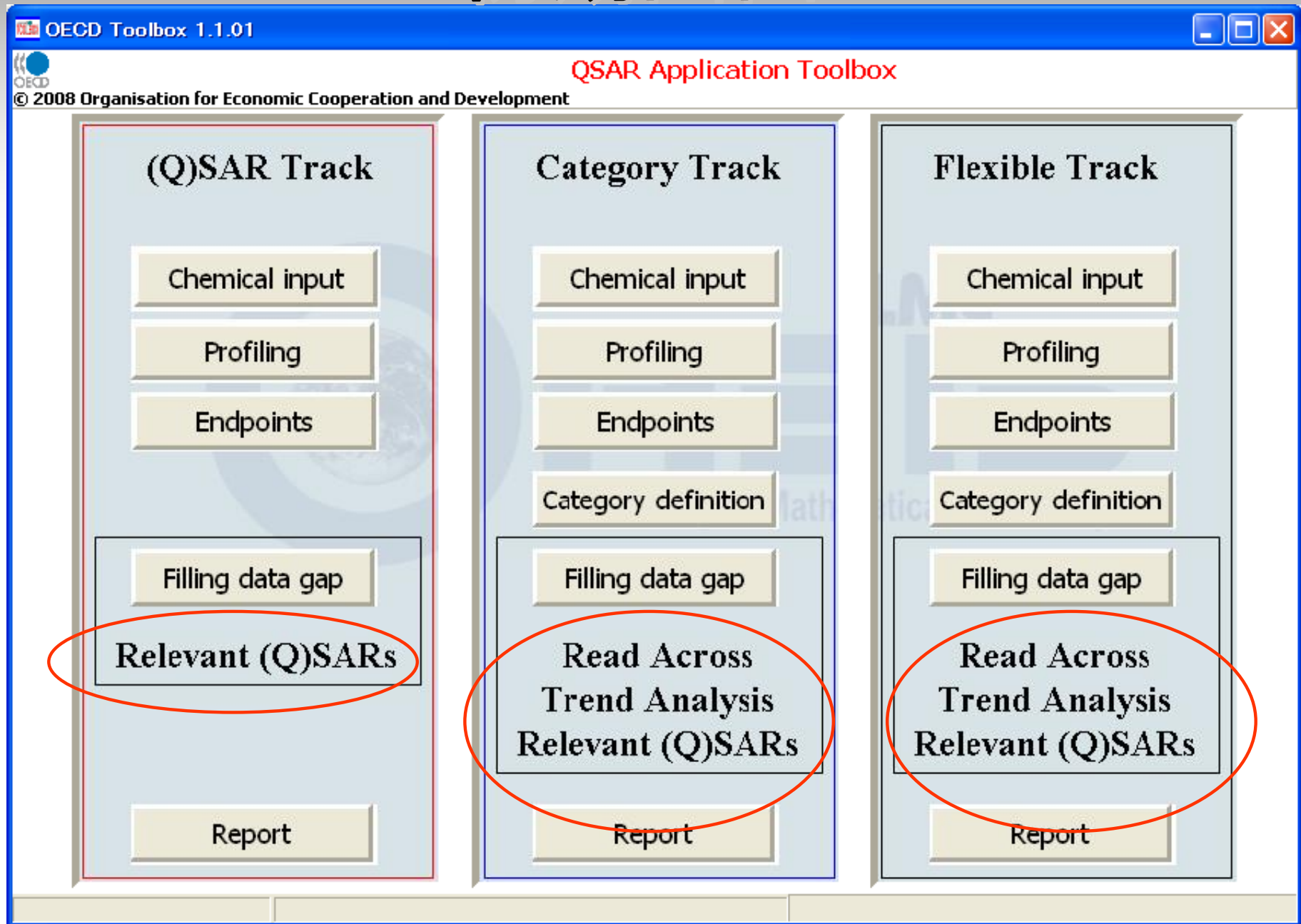
OECD/QSARモデルの規制利用 のためのバリデーション原則

- 1: Defined Endpoint (定義されたエンドポイント)
- 2: Unambiguous Algorithm (曖昧でないアルゴリズム)
- 3: Defined Domain of Applicability (定義された適用可能領域)
- 4: Appropriate Measures of Goodness-of-fit, Robustness and Predictivity (モデルの当てはまりの良さ・頑健さ・予測可能性に関する適切な指標)
- 5: Mechanistic Interpretation, if possible (可能ならば、反応機構の面からの解釈)

OECD (Q)SAR Application Toolbox

- 例えば、以下の事例に用いることが可能
 - Read-Across又はTrend Analysisを用いることで、類似物質の確認及びその実験データの取得、データギャップの補完
 - メカニズム又は作用機序に基づいた、大量の化学物質のカテゴライズ
 - 登録されているQSARモデルを用いたデータギャップの補完
 - Read-Acrossを行なうため、潜在的な類似物質のロバストネスの評価。
 - QSARモデル構築

初期画面



各ステップの機能

- Chemical Input (化学物質の入力)
- Profiling (プロファイリング)
 - ターゲット物質について、官能基の個数やDNA結合性などの属性情報を入手
- Endpoints (エンドポイント)
 - ターゲット物質について、毒性値などのエンドポイントとなる情報を入手
- Category Definition (カテゴリーの定義)
 - ターゲット物質と類似する物質を探索しカテゴリーを定義
 - 類似物質のエンドポイント情報を入手
- Filling Data Gap (データギャップ補完)
 - Read-across、Trend analysis、(Q)SARを用いることで、データギャップを補完
- Report (レポート作成)

KATE, OECD (Q)SAR Application Toolboxデモ

- 百聞は一見にしかず！！
- 下記ページもご参照下さい。
 - 過年度セミナーテキスト ダウンロードページ
<http://www.nies.go.jp/risk/seminar.html>
 - 生態毒性予測システムKATE 関連情報
<http://kate.nies.go.jp/>
 - OECD (Q)SAR Project 関連情報
<http://oecd.org/env/existingchemicals/qsar>
 - 環境リスク研究センター ホームページ
<http://www.nies.go.jp/risk/index.html>



ご静聴ありがとうございました