

生態毒性予測システム「KATE2025」インターネット版 操作マニュアル（2025年3月27日版）



※ KATE2025は、化学物質の生態毒性に関する

- ・ 魚類急性毒性試験における半数致死濃度（LC₅₀）
- ・ ミジンコ急性遊泳阻害試験における半数影響濃度（EC₅₀）
- ・ 藻類生長阻害試験における半数影響濃度（EC₅₀）
- ・ 魚類初期生活段階毒性試験における無影響濃度（NOEC）
- ・ ミジンコ繁殖試験における無影響濃度（NOEC）
- ・ 藻類生長阻害試験における無影響濃度（NOEC）

を予測するシステムです。

※ 本システムで得られた予測結果は、「化学物質の審査及び製造等の規制に関する法律」に基づく届出に必要な生態毒性試験結果として利用することはできません。化学物質の生態毒性影響の程度についての参考としてご利用ください。

ご質問等がございましたら、下記までお問い合わせ下さい。

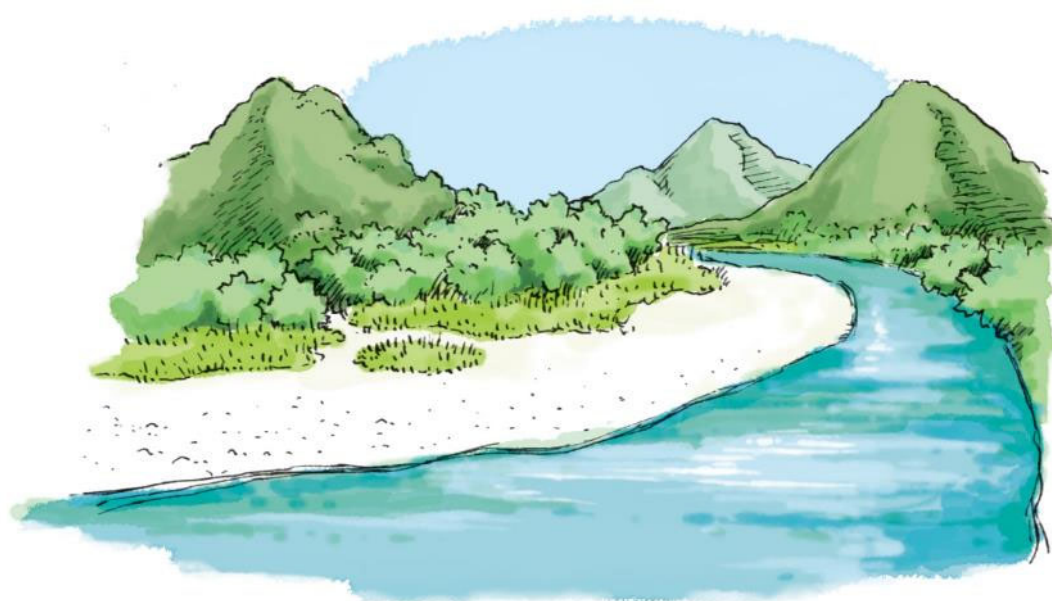
国立研究開発法人 国立環境研究所 環境リスク・健康領域 KATE担当 kate@nies.go.jp

Copyright(C) 2019-2025 Ministry of the Environment, Government of Japan.

All Rights Reserved

マニュアル改訂履歴

バージョン	発行日	改訂履歴
第0.1版	2018年3月29日	KATE2017 on NET β版向けマニュアル
第0.8版	2019年1月30日	KATE2017 on NET 正式版向け暫定版
第0.9版	2019年3月29日	KATE2017 on NET 正式版向け暫定版
第1.0版	2019年5月24日	KATE2017 on NET 正式版対応マニュアル
第1.0.1版	2019年6月4日	説明文を一部修正
第1.0.2版	2019年7月30日	JSME Editorに関する記述を一部変更
	2020年2月3日	KATE2020 (version 1.0) 公開
第2.0版	2020年3月19日	KATE2020 (version 1.0) 向け対応マニュアル
	2021年1月28日	KATE2020 (version 2.0) 公開
第3.0版	2021年1月28日	KATE2020 (version 2.0) 向け対応マニュアル
第4.0版	2022年3月30日	KATE2020 (version 3.0) 向け対応マニュアル
	2022年3月30日	KATE2020 (version 3.0) 公開
第5.0版	2023年3月30日	KATE2020 (version 4.0) 向け対応マニュアル
	2023年3月30日	KATE2020 (version 4.0) 公開
	2023年8月4日	KATE2020 (version 4.1) 公開
第5.0版	2024年3月25日	KATE2020 (version 5.0) 向け対応マニュアル
	2024年3月25日	KATE2020 (version 5.0) 公開
第6.0版	2025年3月27日	KATE2025 (version 1.0) 向け対応マニュアル
	2025年3月27日	KATE2025 (version 1.0) 公開



目次

略語一覧	5
1. はじめに	7
(1) 生態毒性予測システム「KATE」：KASHINHOU TOOL FOR ECOTOXICITYとは	7
(2) KATE2025版とは	7
(3) KATE2020 VERSION 5.0からKATE2025 VERSION 1.0への主な変更点	7
(4) KATE2020 VERSION 4.0からVERSION 5.0への主な変更点	7
(5) KATE2020 VERSION 3.0からVERSION 4.0への主な変更点	8
(6) KATE2020 VERSION 2.0からVERSION 3.0への主な変更点	8
(7) KATE2020 VERSION 1.1からVERSION 2.0への主な変更点	8
(8) KATE2020 VERSION 1.0からKATE2020 VERSION 1.1への主な変更点	8
(9) KATE2017 ON NETからKATE2020 VERSION 1.0への主な変更点	9
(10) KATE2017 ON NETまでの開発経緯	9
(11) サポートケミカルについて	9
(12) LOG Pについて	9
(13) 免責事項	10
(14) 謝辞	10
(15) 参考文献	11
2. KATE2025の概要	12
(1) QSAR予測手順の概要	12
(2) QSAR式の割当、毒性値の計算、及び適用領域判定に関する詳細	14
3. ログイン方法	18
4. 化学物質情報の入力（INPUT画面）	19
(1) 予測できない化学物質について	19
(2) 入力方法1：SMILESの直接入力	19
(3) 入力方法2：描画した構造式からのSMILESへの変換	20
(4) 入力方法3：CAS番号や物質名からのSMILESへの変換	21
(5) 入力LOG Pの利用	22
(6) 入力CAS番号および入力物質名の利用	22
(7) KOWWIN TM 計算のスキップ	22
5. QSAR予測結果の表示（RESULTS画面）	23
(1) 予測対象物質の基本情報	24

(2) QSAR予測結果	25
6. QSARクラス情報の詳細表示 (VERIFY QSAR画面)	28
(1) QSARクラスの基本情報	29
(2) グラフ関連部分	30
(3) 予測対象物質情報.....	33
(4) 回帰式情報	33
(5) 最後にクリックされた物質	34
(6) QSARクラスに含まれる物質情報の表示切り替え機構	35
(7) 構造式一覧 (STRUCTURE VIEWタブ)	35
(8) 物質データ (TABLE VIEWタブ)	37
(9) 構造クラス定義	38
(10) 予測対象物質の部分構造一覧.....	39
7. 複数の化学物質の予測 (UPLOAD 画面)	41
(1) 入力ファイルフォーマットについて.....	41
(2) 予測実行手順.....	42
(3) QSAR予測結果 (複数の化学物質)	44
8. 一括印刷フォーマット表示 (PRINT FORMAT画面)	47



略語一覧

EC₅₀: 50% Effective Concentration (半数影響濃度)

試験水に溶解した化学物質などによって、半数 (50%) の試験生物に対して影響を与えると考えられる濃度

KATE: KAshinoh Tool for Ecotoxicity

国立環境研究所 環境リスク・健康研究センターにおいて研究・開発されている生態毒性QSARシステム。「ケイト」と読む。

KOWWIN™:

US EPAなどが開発しているEPI Suite™ (Estimation Programs Interface: 化学物質を迅速にスクリーニングするためのアプリケーションなどで使用されることを目的としたツール) に含まれる化学物質のlog P推定プログラム。

LC₅₀: 50% Lethal Concentration (半数致死濃度)

試験水に溶解した化学物質などによって、半数 (50%) の試験生物を死亡させる濃度。

log P: The logarithm of the octanol/water partition coefficient

(オクタノール/水分配係数)

ある化学物質について、1-オクタノールと水の2つの溶媒中の平衡状態における濃度比を常用対数で表したもの。化学物質の疎水性を表す指標とされている。

<https://www.eic.or.jp/ecoterm/?act=view&serial=295>

(EICネットの用語解説。2025年3月1日アクセス)

NOEC: No Observed Effect Concentration (無影響濃度)

対照区と比較して統計的に有意な(有害)影響が認められなかった最高濃度であり、LOEC (最小影響濃度) のすぐ下の濃度区である。

<https://www.env.go.jp/content/000212303.pdf>

<https://www.env.go.jp/chemi/report/ierac22/index.html>

(環境省 化学物質の環境リスク評価 第2巻 第1編 参考2 用語集等より。2025年3月1日アクセス)

QSAR: Quantitative Structure-Activity Relationships (定量的構造活性相関)

化学物質の構造上の特徴又は物理化学定数と生物学的活性(毒性等)の相関関係を構造活性相関(SAR: Structure-Activity Relationship)といい、定量的なものを定量的構造活性相関(QSAR: Quantitative Structure-Activity Relationship)という。両者を併せて(Q)SARと記載することもある。例えば、化学物質の特徴的な部分構造の有無やlog Kow (1-オクタノール/水分配係数)と化学物質の毒性等との関係「QSAR式」を求め、QSAR式を利用して毒性等を推測する。部分構造を手掛かりに化学物質を分類したものをQSARクラスという。QSARクラスごとに構築したQSAR式を用いて毒性等を定量的に算出する仕組みをQSARモデルと呼ぶ。

<https://www.env.go.jp/content/000212303.pdf>

<https://www.env.go.jp/chemi/report/ierac22/index.html>

(環境省 化学物質の環境リスク評価 第2巻 第1編 参考2 用語集等より。2025年3月1日アクセス)



SMARTS: SMiles ARbitrary Target Specification

SMILESを拡張した部分構造を表現するための識別子。

https://www.daylight.com/dayhtml_tutorials/languages/smarts/

(Daylight社詳細解説。2025年3月1日アクセス)

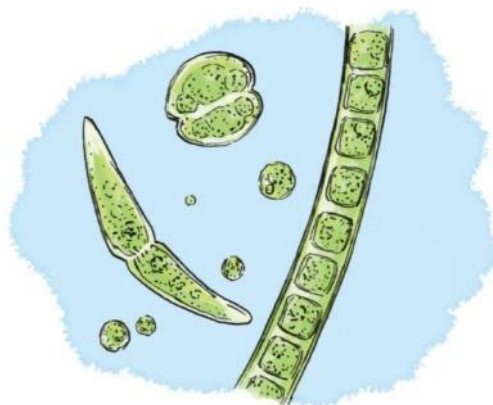
SMILES: Simplified Molecular Input Line Entry System

化合物の分子構造等を印刷可能な文字で線形表記した識別子。

<https://www.daylight.com/smiles/index.html>

(Daylight社詳細解説。2025年03月1日アクセス)

US EPA: United States Environmental Protection Agency (米国環境保護庁)



1. はじめに

(1) 生態毒性予測システム「KATE」：KAshinhou Tool for Ecotoxicityとは

国立研究開発法人 国立環境研究所 環境リスク・健康領域において、環境省の請負業務により研究・開発された生態毒性 QSAR システムです¹⁾。

現在の最新バージョンである KATE2025 (version 1.0)では、化学物質の構造から以下の毒性値の種類²⁾に対する予測結果が得られます。

- ・魚類急性毒性試験 (OECD TG 203) における半数致死濃度 (LC₅₀)
- ・魚類初期生活段階毒性試験 (OECD TG 210) における無影響濃度 (NOEC)
- ・ミジンコ急性遊泳阻害試験 (OECD TG 202) における半数影響濃度 (EC₅₀)
- ・ミジンコ繁殖試験 (OECD TG 211) における無影響濃度 (NOEC)
- ・藻類生長阻害試験 (OECD TG 201) における半数影響濃度 (EC₅₀) 及び無影響濃度 (NOEC)

化学物質の CAS 番号検索や構造式エディタを用いた作図等を用いて SMILES 記法による入力を行い、log P を記述子とした QSAR 予測を行います。

KATE2025 の構築に当たっては、環境省が実施した生態毒性試験結果 (魚類急性毒性試験、ミジンコ急性遊泳阻害試験、魚類初期生活段階毒性試験、ミジンコ繁殖試験、藻類生長阻害試験)³⁾及び US EPA のファットヘッドミノール・データベースの魚類急性毒性試験結果⁴⁾をトレーニングセットデータとして用いています。試験結果が追加された場合には QSAR クラスの回帰式の再計算を行っています。

(2) KATE2025版とは

2020年2月から公開を行っていた KATE2020 (version 1.0~5.1) に改良を加え、2025年3月に KATE2025 version 1.0 へメジャーアップデートを行いました。インターネット上のブラウザ画面で操作を行います。

(3) KATE2020 version 5.0からKATE2025 version 1.0への主な変更点

ユーザーインターフェースの改良

- ① 複数予測の場合に、一回の計算で 1000 物質まで予測できるように変更
- ② 予測結果の表のフィルタリング (毒性値と統計値による絞り込み) 機能の簡素化
- ③ QSAR クラス情報の詳細表示画面における物質情報の表示をタブ化
- ④ その他、全体的な表示や操作性の修正

QSAR モデルの更新

- ① 魚類及びミジンコの慢性に関する生態影響試験結果の見直しを行い、毒性値の再計算が必要なものは最新のテストガイドラインに基づいて再計算を実施し、異性体混合物等はサポートケミカルに変更
- ② 2023 年度に生態影響試験が実施された硫黄含有化学物質 2 物質をトレーニングセットデータに追加
- ③ ①②の結果を受けて QSAR クラスの回帰式の再計算を実施

(4) KATE2020 version 4.0からversion 5.0への主な変更点

ユーザーインターフェースの改良

- ① 入力画面および複数物質予測結果画面の改良
- ② 複数物質予測結果のダウンロード機能の追加
- ③ 構造クラス定義の表示方法の改良

QSAR モデルの更新

- ① 藻類の急性および慢性に関する生態影響試験結果の見直しを行い、毒性値の再計算が必要なものは最新のテストガイドラインに基づいて再計算を実施し、異性体混合物等はサポートケミカルに変更



- ② 2022 年度に生態影響試験が実施された脂肪族一級アミン 2 物質とエポキシド 1 物質をトレーニングセットデータに追加
- ③ ①②の結果を受けて QSAR クラスの回帰式の再計算を実施

(5) KATE2020 version 3.0からversion 4.0への主な変更点

ユーザーインターフェースの改良

- ① ログイン方法を変更し、ユーザ ID とパスワードを使用せずに KATE2020 にログインできるように変更
- ② 予測結果画面や QSAR クラス詳細画面等の修正

QSAR モデルの更新

- ① 魚類急性に関する生態影響試験結果の見直しを行い、毒性値の再計算が必要なものは最新のテストガイドラインに基づいて再計算を実施し、トレーニングセットとしての使用が不適切なものはサポートケミカルに変更
- ② 2021 年度に生態影響試験が実施されたチオール 2 物質（魚類急性、甲殻類急性）とイミド 1 物質（甲殻類急性）をトレーニングセットデータに追加
- ③ ①②の結果を受けて QSAR クラスの回帰式の再計算を実施

QSAR クラス名の更新

- ① すべての QSAR クラスの名称を精査し、一部の QSAR クラスについて名称を変更

(6) KATE2020 version 2.0からversion 3.0への主な変更点

QSAR モデルの更新

- ① 甲殻類急性に関する生態影響試験結果の見直しを行い、毒性値の再計算が必要なものは最新のテストガイドラインに基づいて再計算を実施し、トレーニングセットとしての使用が不適切なものはサポートケミカルに変更
- ② ①の結果を受けて QSAR クラスの回帰式の再計算を実施

QSAR クラス名の更新

- ① 統計指標の基準 ($R^2 \geq 0.7$ and $Q^2 \geq 0.5$ and $n \geq 5$) を満たすすべての QSAR クラスの名称を精査し、一部の QSAR クラスについて名称を変更

(7) KATE2020 version 1.1からversion 2.0への主な変更点

QSAR モデルの更新

- ① デフォルトで表示する QSAR クラスの基準を「 $R^2 \geq 0.7$ and $Q^2 \geq 0.6$ and $n \geq 5$ 」から「 $R^2 \geq 0.7$ and $Q^2 \geq 0.5$ and $n \geq 5$ 」に変更
- ② 毒性値 1 つを修正（転記ミスによる）。このことに付随して、藻類慢性の QSAR クラス「CNOS_X basic aromatic n unreactive」の QSAR 式を変更

表示・操作の改良

- ① 部分構造に対する構造判定の追加
- ② 印刷フォーマット表示の追加
- ③ 複数化学物質予測でエラーがあった場合に途中で止まる問題の修正

(8) KATE2020 version 1.0からKATE2020 version 1.1への主な変更点

- ① 予測毒性値が 10^6 以上もしくは 10^{-5} [mg/L]未満の場合は指数表記（例： $2.3e-7$ ）で表示するように変更
- ② 予測毒性値を有効数字 2 桁で表示するように変更
- ③ 化学物質入力の際に、正しい情報を入れたときにもエラーになる場合がある問題の修正



(9) KATE2017 on NETからKATE2020 version 1.0への主な変更点

QSAR モデルの更新

- ① log P の推定方法を ClogP から KOWWIN™ に変更、および変更に伴う一部の QSAR モデルの修正
- ② 予測対象物質が使用する log P 値の優先順位を 1. ユーザ入力値、2. KOWWIN™ 推定値に変更（以前は 1. ユーザ入力値、2. KOWWIN™ 実測値、3. KOWWIN™ 推定値であった）
- ③ log P>6.0 のデータは全て QSAR モデルから除外
- ④ トレーニングセットと QSAR クラスの追加・削除
- ⑤ QSAR 式構築に使用されない物質（サポートケミカル）を参考情報として表示するように変更

表示・操作の改良

- ① QSAR クラスに含まれるトレーニングセットおよびサポートケミカルのデータ一覧表の追加
- ② QSAR クラスに対応する構造クラスの定義一覧の追加
- ③ QSAR 式詳細画面の構造式一覧にソート機能の追加
- ④ QSAR 式詳細画面中グラフにおいて、トレーニングセットの一部を除外して計算される回帰直線と予測区間・信頼区間を連動させた
- ⑤ KOWWIN™ 計算をスキップする機能の追加

(10) KATE2017 on NETまでの開発経緯

現在もホームページ上で公開している KATE2011 版については、2008 年 1 月に試用版、2009 年 3 月に KATE2009 をインターネットで公開しました⁵⁾。2011 年 3 月には、更新版 KATE2011 を公開し、トレーニングセットデータの追加、部分構造の分類ルールの修正、構造判定の変更と皮膚感作性に関する部分構造の追加等を行いました。

その後、現在の KATE2020 版と同じ系列の新しい生態毒性予測システムの開発を開始し、2018 年 3 月に KATE2017 on NET β 版、2019 年 1 月に KATE2017 on NET 正式版 (version 1.0) を公開しました。KATE2011 から KATE2017 への変更では主に、部分構造検索方式の FITS（KATE2011 で使用される部分構造表記法および検索プログラム）から SMARTS 記法および CDK を利用した検索プログラムへの変更、藻類や慢性の毒性値予測の追加、不等号付き毒性値データの追加、構造クラスの導入、部分構造・QSAR クラスの大幅な変更、log P 計算モジュールの ClogP から KOWWIN™ への変更、表示・操作の改良、英語化等を行いました。

(11) サポートケミカルについて

KATE2020 から、以下のデータは QSAR モデル構築には使用せず、サポートケミカル（参考情報）として表示のみを行うようにしました。

- ① log P 推定値>6.0 の化学物質データ
 - ② 不等号付きデータ
 - ③ 異性体混合物等
- なお、②の不等号付きデータは、log P の適用領域内にある場合、構造に関する適用領域の判定に使用しています。

(12) log Pについて

本システムは、化学物質の毒性を予測する際に使用する log P として、US EPA が著作権を有する log P 予測モデル KOWWIN™ を US EPA の許諾を得て使用しています⁶⁾。利用者は下記に示す KOWWIN™ 使用許諾条件について遵守してください。

KOWWIN v1.69 (April 2015)

c 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.

(13) 免責事項

KATE の予測結果は十分な予測精度を保証できるものではありません。本システムは、化学物質の生態毒性影響の程度についての参考情報を得るためのツールの一つとしてご利用ください。環境省および国立環境研究所は KATE による予測毒性値を保証するものではなく、また、KATE による予測毒性値の使用により生じた損害については一切の責任を負いません。

また、現時点では KATE による毒性予測結果を「化学物質の審査及び製造等の規制に関する法律」（化審法）に基づく届出に必要な生態毒性試験結果に代替するものとして利用することはできません。

著作権、リンク等については、KATE ウェブサイト内のサイトポリシーをご覧ください。<https://kate.nies.go.jp/spolicy-ja.html>

(14) 謝辞

KATE2025 は下記のソフトウェアまたはライブラリからの結果を使用させていただいております。ここに記して謝意を表します。

- Open Babel
 - <https://openbabel.org/>
- JSME Molecular Editor
 - <https://jsme-editor.github.io/>
 - B Bienfait and P Ertl, JSME: A free molecule editor in JavaScript, J. Cheminform. 5:24 (2013). doi:10.1186/1758-2946-5-24.
- CDK (Chemistry Development Kit)
 - <https://cdk.github.io/>
 - E Willighagen et al., The Chemistry Development Kit (CDK) v2.0: Atom typing, depiction, molecular formulas, and substructure searching, J. Cheminform. 9:33 (2017). doi:10.1186/s13321-017-0220-4.
 - JW May and C Steinbeck, Efficient ring perception for the Chemistry Development Kit, J. Cheminform. 6:3 (2014). doi:10.1186/1758-2946-6-3.
 - C Steinbeck et al., Recent developments of the Chemistry Development Kit (CDK) - an open-source Java library for chemo- and bioinformatics, Curr. Pharm. Des 12:2111-2120 (2006). doi:10.2174/138161206777585274.
 - C Steinbeck et al., The Chemistry Development Kit (CDK): An open-source Java library for chemo- and bioinformatics, J. Chem. Inf. Comput. Sci. 43:493-500 (2003). doi:10.1021/ci025584y.
- KOWWIN™ (included in EPI Suite™)
 - <https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface>

(上記 URL は全て、2025 年 3 月 1 日アクセス)



(15) 参考文献

- 1) <https://kate.nies.go.jp> (2025年3月1日アクセス)
- 2) <https://www.env.go.jp/chemi/sesaku/01.html> (2025年3月1日アクセス)
- 3) <https://www.env.go.jp/chemi/sesaku/seitai.html> (2025年3月1日アクセス)
- 4) https://archive.epa.gov/med/med_archive_03/web/html/fathead_minnow.html
(2025年3月1日アクセス)
- 5) A. Furuhashi, T. Toida, N. Nishikawa, Y. Aoki, Y. Yoshioka, and H. Shiraishi:
Development of an ecotoxicity QSAR model for the KAsinhou Tool for Ecotoxicity
(KATE) system, March 2009 version, SAR QSAR Environ. Res., 21 (5), 403 (2010).
- 6) <https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface>
(2025年3月1日アクセス)



2. KATE2025の概要

(1) QSAR予測手順の概要

KATE2025は、化学物質の毒性予測を下記の流れで行います（図2-1）。

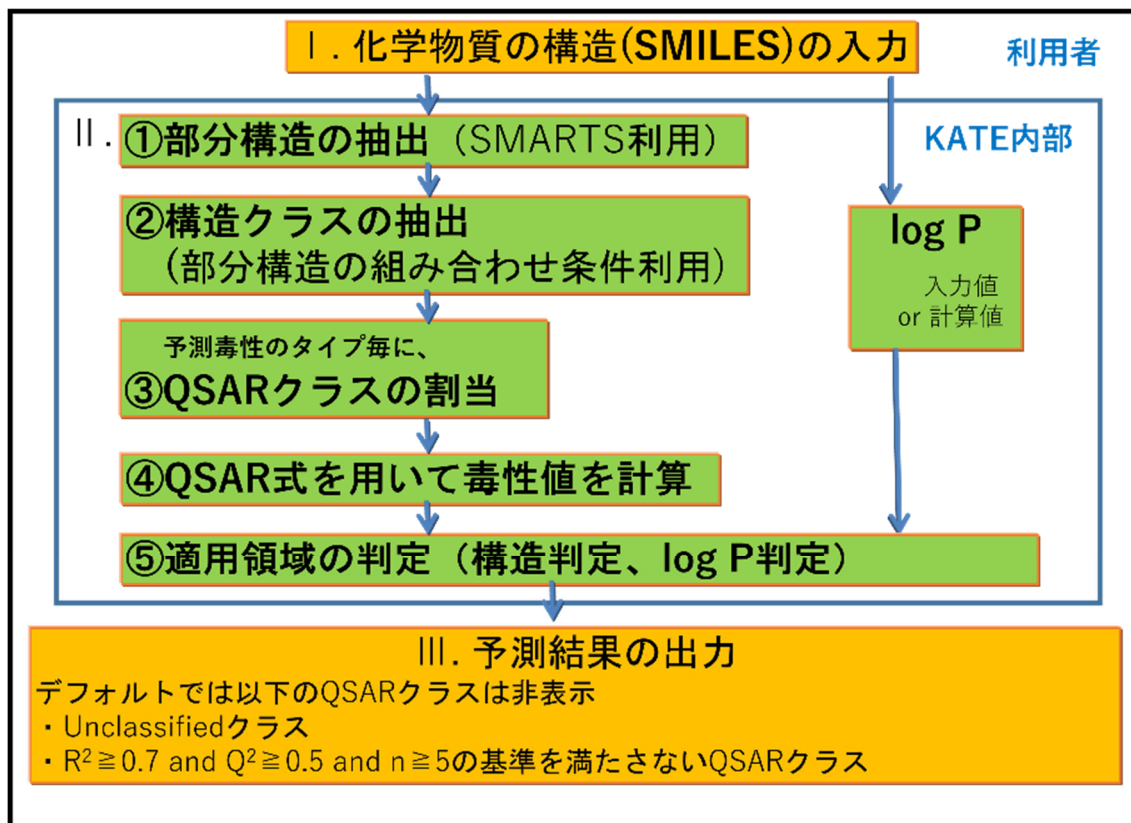


図2-1 KATE2025のQSAR予測フロー

I. 利用者による化学物質の構造 (SMILES記法による) の入力

II. QSAR式、毒性値の決定、及び適用領域判定に関する処理

- ① 予測対象物質 (入力された化学物質) の部分構造を抽出
- ② 複数の部分構造の組み合わせによる構造クラス*1を抽出
- ③ 予測毒性のタイプごとに、構造クラスに対応するQSARクラス*2を割当 (複数のクラスに割当てられることもあります。)
- ④ 割り当てられたQSARクラスごとに、QSAR式*3を用いて毒性値を計算
- ⑤ 適用領域の判定 (構造判定とlog P判定)

*1 各部分構造の個数条件のAND/ORによる組み合わせにより定義した分類 (38ページの「構造クラス定義」参照)

*2 各予測毒性のタイプでの物質の構造に基づいて定義した分類

*3 QSARクラスに含まれるトレーニングセットで形成されたモデル。ここではlog Pを記述子とする単回帰式



III. 予測結果の出力

- ① 予測対象物質が何らかの予測毒性のタイプでどのQSARクラスにも分類されなかった場合、Unclassifiedクラス*4に割当てられます。
 - ② デフォルトではUnclassifiedクラス、及び統計指標の基準 ($R^2 \geq 0.7$ and $Q^2 \geq 0.5$ and $n \geq 5$) を満たさないQSARクラス*5は非表示にしています。
- *4 どのQSARクラスにも分類されなかった場合に割当てられるQSARクラス
 *5 R^2 , Q^2 , n は、各QSARクラスに対してあらかじめ計算されています。

具体的な例として、化学物質 1-pyridin-3-ylethanone を予測したときの予測フローを図2-2に示します。

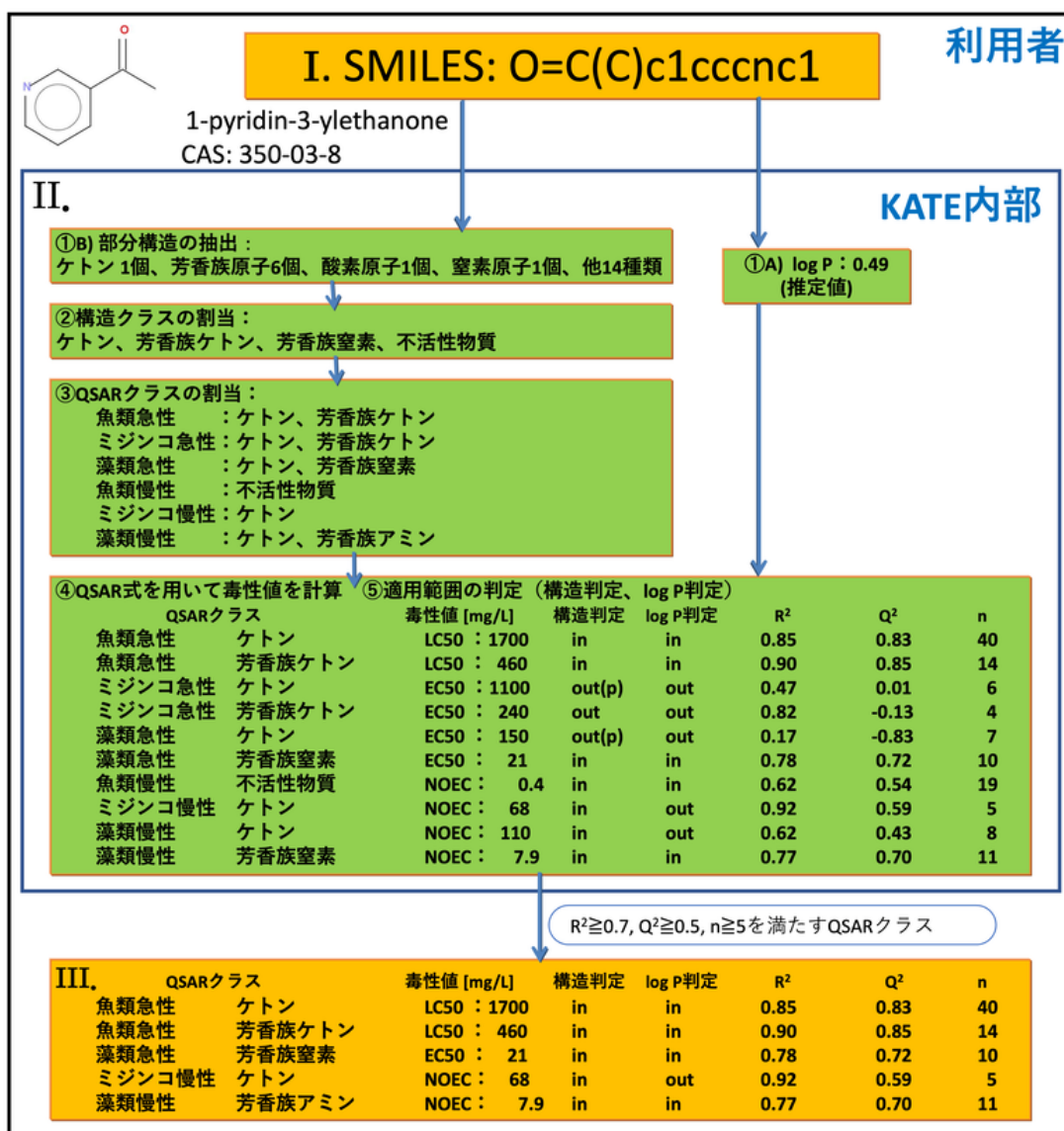


図2-2 KATE2025のQSAR予測フロー(1-pyridin-3-ylethanoneの予測例)

※ 図2-2中のケトン、芳香族ケトン、芳香族窒素、低反応性物質のKATE2025における実際の名称は以下のようになっています。

ケトン : COS_X ketone unreactive



芳香族ケトン	: COS_X ketone unreactive aromatic
芳香族窒素	: aromatic n unreactive Alga
低反応性物質	: CNO_X unreactive Fish Chronic, w/ N,O

(2) QSAR式の割当、毒性値の計算、及び適用領域判定に関する詳細

KATE2025 では、利用者が入力した化学物質の構造（SMILES 記法で作成）に基づき、以下の処理を行い QSAR 式及び毒性値を予測します。以下に、簡単な説明を行います。

① 部分構造の抽出

KATEで定義されている部分構造一覧（SMARTS記法で作成）をもとに、予測対象物質に含まれる各部分構造の個数を計算します。SMARTSを利用した部分構造個数計算にはCDKライブラリを利用しています。

② 構造クラスの割当

構造クラス定義一覧をもとに、予測対象物質の構造に合致する全ての構造クラスを抽出します。

③ QSARクラスの割当

各予測毒性のタイプについて、QSARクラス定義一覧をもとに予測対象物質の構造クラスに対応するQSARクラスを割り当てます。KATE2025では、同じ予測毒性タイプに対して複数のQSARクラスが割り当てられる場合があります。また、どのQSARクラスにも分類されなかった場合は、Unclassified クラスに割り当てられ、毒性の予測が行われません。

④ QSAR式による毒性値の計算

各QSARクラスにはQSAR式が割り当てられており、予測対象物質のlog Pの値をQSAR式に代入することにより、 $\log(1/\text{毒性値}[\text{mmol/L}])$ を計算します。次に、予測対象物質の分子量を用いて毒性値[mg/L]に単位変換します。

⑤ 適用領域の判定

KATE2025では、予測対象物質の毒性値が、予測結果として適用できる範囲内にあるかどうかを判定します。A) 構造による判定とB) log Pによる判定の2つを行い、両方とも適用領域内の場合に、毒性値が適用領域と判定されます。

A) 構造判定

KATE2025では、「構造判定用部分構造」*1の比較により、予測対象物質の構造が当該QSARクラスの適用領域内であるかどうかについて判定します（図2-3）。判定には下記の3つの場合があり、「in」又は「in(p)」の場合に当該QSARクラスを、構造に関して適用領域内と判定しています。

in : 適用領域内

予測対象物質に含まれる「構造判定用部分構造」の全てが、当該 QSAR クラスの「構造判定用部分構造リスト*2」に含まれる場合（図2-3におけるピンクとオレンジの範囲）、または予測対象物質が持つ部分構造に「構造判定用部分構造」が1つも含まれていなかった場合。

in(p) : 条件付き適用領域内

「in」の条件には合致しないが、予測対象物質の「構造判定用部分構造」全てが、当該 QSAR クラスの「構造判定用部分構造リスト」、あるいは Narcotic Group*3 クラスの「構造判定用部分構造リスト」に含まれる場合（図2-3におけるピンクとオレンジとイエローの部分）。

out : 適用領域外



「in」と「in(p)」の何れの条件にも合致しない場合。すなわち、予測対象物質の「構造判定用部分構造」に、当該 QSAR クラスの「構造判定用部分構造リスト」と Narcotic Group クラスの「構造判定用部分構造リスト」の何れにも含まれない部分構造がある場合。（図 2-3 においてグレー部分の構造が含まれる場合）

*1 KATE で定義されている部分構造一覧の中で、構造判定にも使用されるもの。KATE2025 では 175 個の部分構造が該当する（詳細は KATE2025 技術文書*参照）。

※ https://kate.nies.go.jp/doc/KATE2025_tech_doc-ja.pdf

*2 当該 QSAR クラスに含まれる、トレーニングセットおよび log P 判定が「in（適用領域内）」である不等号付きデータの物質が持つ、部分構造のリスト。

*3 特異的な生理活性作用に基づかないベースライン毒性（麻酔作用）。KATE2025 では、脂肪族炭化水素、スルホキド、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 QSAR クラスが予測毒性のタイプ毎に用意されており、これらをまとめた QSAR クラスを Narcotic Group として各予測毒性のタイプで再定義しています。

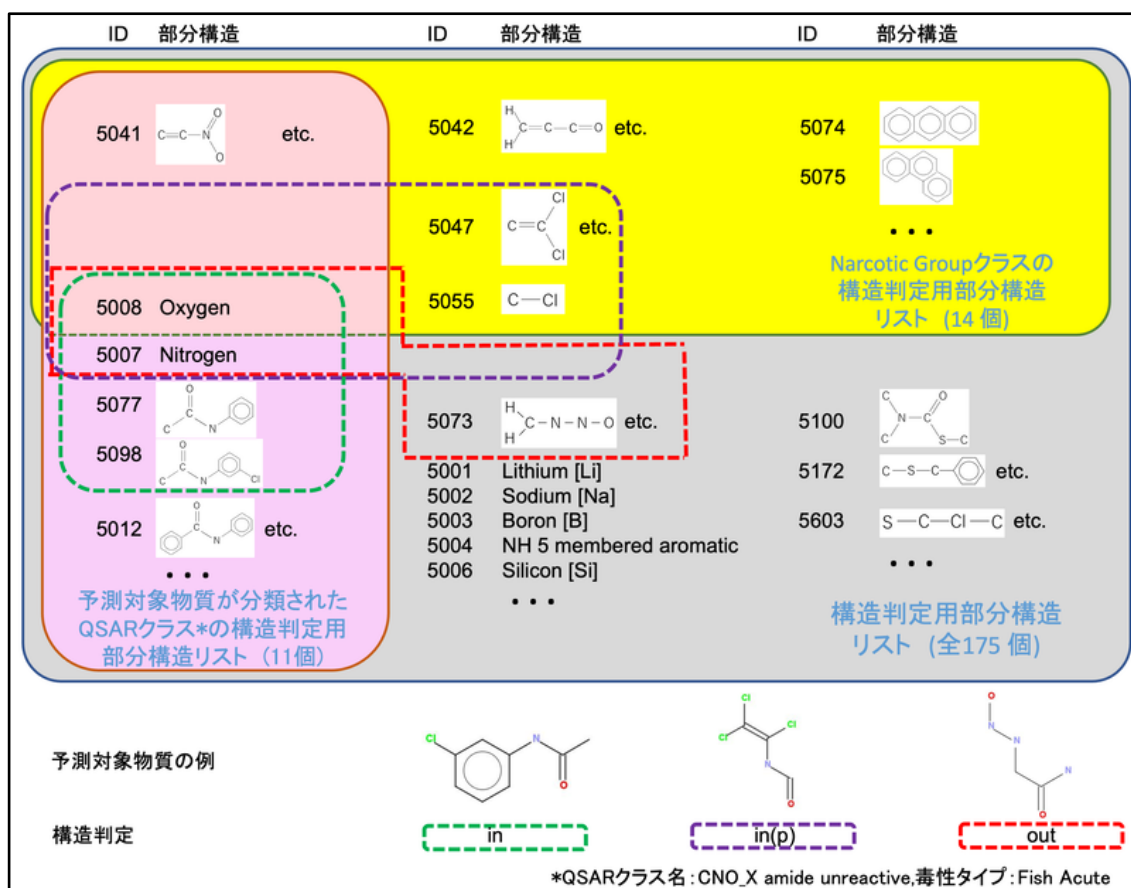


図 2-3 構造判定の例

※ 緑枠は「予測対象物質の例」の一番左の物質で予測して緑枠の部分構造が抽出されることにより構造判定が緑枠の「in」となることを示しています。紫枠と赤枠も同様です。



※ 「etc.」が付くものは、同じIDでも複数の部分構造があり、1つの例のみ示しています。

B) log P判定

KATE2025では、予測対象物質のlog P値が当該QSARクラスのトレーニングセットのlog Pの最小値と最大値の間にあるかどうかで適用領域内にあるかどうかを判定します。なお、高疎水性のため予測精度が低いと考えられる $\log P > 6$ の物質は全て適用領域外としています（KATE2020版における変更点参照）。

in：適用領域内（図2-4）

out：適用領域外。ただし、下記の out(p) になる場合は除く（図2-5）。

out(p)：適用領域外。ただし、予測対象物質の log P 値は当該 QSAR クラスの参考情報であるトレーニングセットとサポートケミカルを併せた全物質の log P の最小値と最大値の内側に存在する（図2-6）。

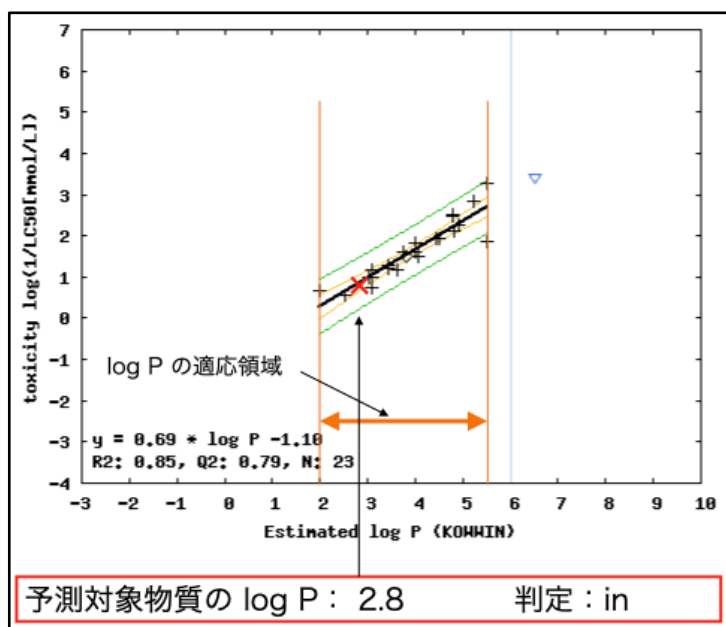


図2-4 log P判定の例 (in)



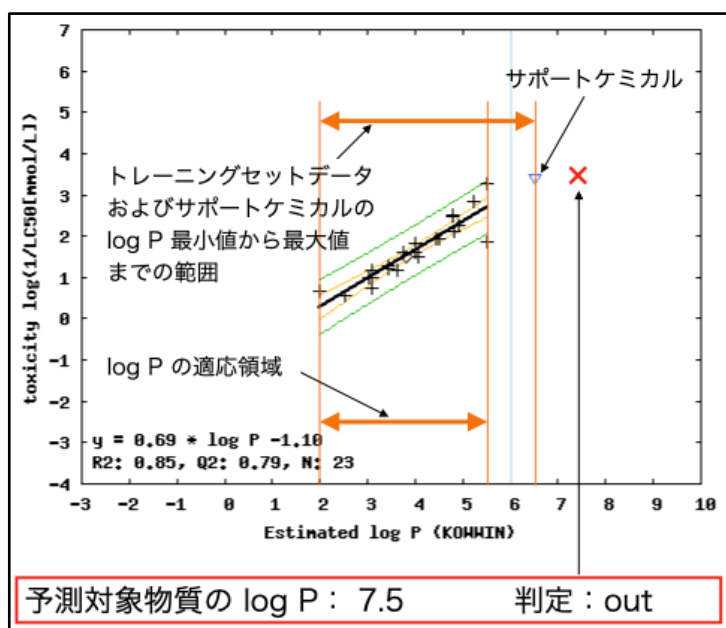


図 2 - 5 $\log P$ 判定の例 (out)

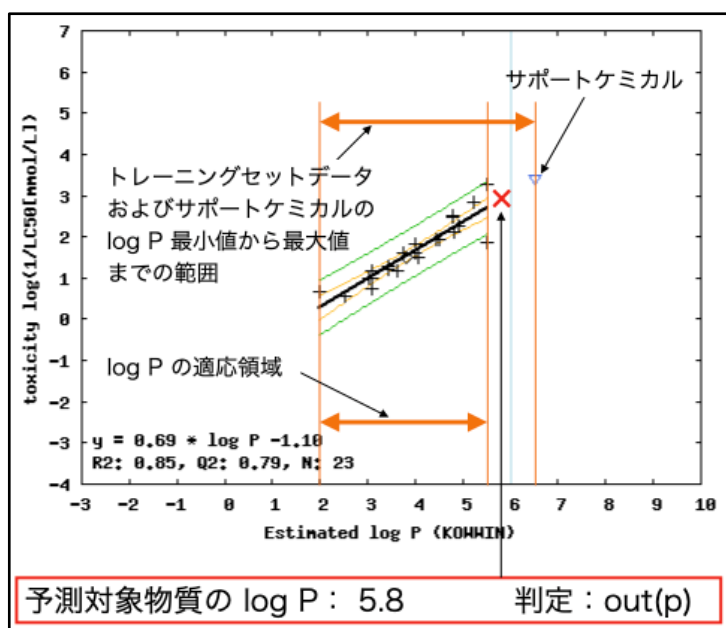


図 2 - 6 $\log P$ 判定の例 (out(p))



3. ログイン方法

2023年3月にリリースされたKATE2020 version 4.0より、ログインするためのユーザIDとパスワードは不要になりました。ユーザは図3に示すKATE2025のログイン画面 (<https://kate.nies.go.jp/nies/index.php>) にアクセスし、(1)免責事項に同意するチェックボックスにチェックを入れ、(2)「Start the session」 ボタンをクリックするとログインすることができます。SMILES等のユーザが入力したデータおよび出力結果（予測毒性値やQSARクラス等）はKATE2025サーバではなくセッションにのみ保持され、セッションが切れると自動的にそれらは削除されます。セッションはウェブブラウザを閉じたりKATE2025の操作を一時間停止したりすると自動的に切れます。また、予測を行う度に過去の情報は上書きされ、削除されます。

KAshinhou Tool for Ecotoxicity
KATE2025 version 1.0

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) → I agree to and accept the terms of agreement above.

(2) →

図3 ログイン画面



4. 化学物質情報の入力（Input画面）

KATE2025にログインするとFront Page が表れ、上段に以下の画面（Input画面）が表示されます（図4-1）。

Front Page

Input SMILES of Your Chemical

READ ME FIRST | Get information using Chemical Identifier Resolver or Generate SMILES using JSME Editor

SMILES (Required): Predict

CAS RN:

Chemical Name: Clear

log P:

Skip KOWWIN: · When any error occurs in log P estimation by KOWWIN, you can skip it.

Output from <https://cactus.nci.nih.gov> may be shown here.

▶ about SMILES notation in KATE2025
▶ about log P

図4-1 化学物質情報の入力画面

KATE2025における予測は、SMILESに基づいて行います。化学物質のSMILESを入力するには以下の方法があります。

入力方法1：SMILESを直接入力する方法

入力方法2：描画した構造式をSMILESに変換して入力する方法

入力方法3：CAS番号や物質名からSMILESを取得して入力する方法

(1) 予測できない化学物質について



注意!

KATEにおいて予測対象となる化学物質は、基本的には有機化合物です。

例外として一部の無機窒素化合物（ヒドラジン等）も含まれます。

KATE2025では、以下i)~v)に該当する化学物質は予測できません。

- i) C、Nの何れの元素も含まない SMILES
- ii) C、N、O、F、Si、P、S、Cl、As、Br、Sn、I以外の元素を含む SMILES
- iii) イオンを含む SMILES（ただし、ammonium [N+] と [n+] は入力可能）
- iv) 混合物を表す（記号「.」を含む） SMILES
- v) [Na], [K], [Li], [Na+], [K+], [Li+]等を含む SMILES の場合、プロトン化した形式に置き換える必要があります。例えば、「c1ccccc1O[Na]」は「c1ccccc1O」に置き換える必要があります。

入力 SMILES が上記 i)~v) に該当する場合、エラーメッセージ画面が表示されます。

(2) 入力方法1：SMILESの直接入力

ログイン後、前ページ図4-1の画面が表示されるので、赤枠で囲んだ入力ボックスに、毒性予測を行いたい化学物質のSMILESを入力します。

SMILESを入力したら、予測（Predict）ボタンをクリックします（図4-2）。この操作を行うとQSAR予測結果画面に進みます。ここでは予測対象物質の例として、NCc1ccccc1を入力します。



図 4 - 2 SMILES 入力ボックス横の予測 (Predict) ボタン

(3) 入力方法 2 : 描画した構造式からの SMILES への変換

JSME Molecular Editor (構造式エディタ、以下、及びウェブ上では JSME Editor と表記します) を利用して、毒性予測を行いたい化学物質の構造式を描画し、それを SMILES に変換することができます。

図 4 - 3 の赤枠の「Generate SMILES using JSME Editor」をクリックします。

図 4 - 3 「JSME Editor」を起動するボタン

JSME Editor (図 4 - 4) が起動するので、構造式 (ここではフェノールの例を示します) を描画し、「Submit smiles to KATE」ボタンをクリックします。JSME Editor については、JSME Editor のウェブページを参照ください。(<https://jsme-editor.github.io/>, 2025 年 3 月 1 日アクセス)

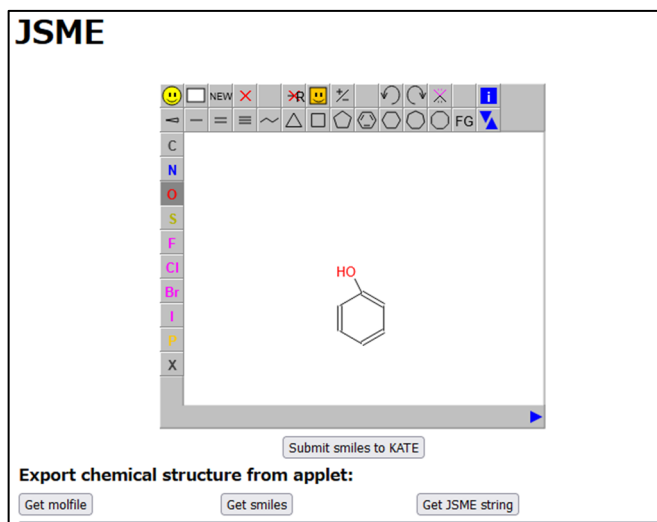


図 4 - 4 JSME Editor の画面

クリックすると、描画した構造式が SMILES に変換され、化学物質の入力画面の SMILES 入力ボックスに入力されます (図 4 - 5)。その後、予測 (Predict) ボタンをクリックすると予測結果画面に進みます。



READ ME FIRST Get information using Chemical Identifier Resolver or Generate SMILES using JSME Editor

SMILES (Required): Predict

図 4-5 構造式から SMILES に変換した結果

(4) 入力方法 3 : CAS番号や物質名からの SMILES への変換

「Get information using Chemical Identifier Resolver」ボタンを使用して、CAS番号や物質名から SMILES を取得することが出来ます。以下の手順では CAS 番号から SMILES を取得する例を示します。Chemical Identifier Resolver については、Chemical Identifier Resolver のウェブページを参照ください。<https://cactus.nci.nih.gov/chemical/structure>, 2025年3月1日アクセス)

毒性予測を行いたい化学物質の CAS 番号を CAS 入力ボックスに入力し、「Get information using Chemical Identifier Resolver」ボタンをクリックします (図 4-6)。

READ ME FIRST Get information using Chemical Identifier Resolver or Generate SMILES using JSME Editor

SMILES (Required):

CAS RN#:

Chemical Name:

図 4-6 CAS 番号入力ボックスと「Get information using Chemical Identifier Resolver」ボタン

その後、CAS 番号から SMILES が取得され、SMILES 入力ボックスに入力されます (対応する構造式も出力され、取得された物質名 (IUPAC 名) も Name ボックスに入力されます。図 4-7)。この後「Predict」ボタンをクリックすると予測結果画面に進みます。

READ ME FIRST Get information using Chemical Identifier Resolver or Generate SMILES using JSME Editor

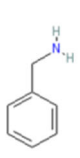
SMILES (Required): Predict

CAS RN#:

Chemical Name: Clear

log P:

Skip KOWWIN™: · When any error occurs in log P estimation by KOWWIN™, you can skip it.



Information obtained from CAS RN.

SMILES: NCc1ccccc1
CAS RN#: 100-46-9
IUPAC NAME: phenylmethanamine

図 4-7 CAS 番号から SMILES と構造式に変換した結果



(5) 入力log Pの利用

予測対象物質のlog P値として、入力値を指定することが出来ます（図4-8）。入力は任意です。入力したlog P値が毒性予測の際に優先的に利用されます。予測対象物質のlog P値が明らかな場合などにご利用ください。

READ ME FIRST Get information using Chemical Identifier Resolver or Generate SMILES using JSME Editor

SMILES (Required): Predict

CAS RN®:

Chemical Name: Clear

log P:

Skip KOWWIN™: · When any error occurs in log P estimation by KOWWIN™, you can skip it.

図4-8 log Pの入力

(6) 入力CAS番号および入力物質名の利用

CAS番号や物質名も、ユーザが任意で指定することができます（図4-9）。入力は必須ではありません。入力した場合は、結果画面には入力した情報がそのまま表示されます。

SMILES (Required):

CAS RN®:

Chemical Name:

log P:

Skip KOWWIN™: · When any error occurs in log P estimation by KOWWIN™, you can skip it.

図4-9 CAS番号と物質名の入力

(7) KOWWIN™計算のスキップ

KOWWIN™に対応していないSMILESが入力されることにより、KOWWIN™によるlog P計算中にエラーになることがあります。この場合には、「Skip KOWWIN Calculation」のチェックボックスにチェックをしてPredictボタンをクリックすることで、KOWWIN™によるlog P計算をスキップすることが出来ます（図4-10）。

log P計算をスキップした場合、予測毒性値を計算するにはlog Pのユーザ入力値が必要となります。log Pを入力しない場合でもQSARクラス分類は行われますが、予測毒性値は計算されません。スキップ後に、予測結果画面の中でlog P入力値を入れて予測毒性値を計算することも可能です。

SMILES (Required):

CAS RN®:

Chemical Name:

log P:

Skip KOWWIN™: · When any error occurs in log P estimation by KOWWIN™, you can skip it.

図4-10 KOWWIN™計算のスキップ



5. QSAR予測結果の表示（Results画面）



予測（Predict）ボタンをクリックした後、計算が終了すると、予測結果画面が表示されます（図5-1）。予測結果画面では、予測対象物質の基本的な情報（CAS番号、物質名、SMILES、分子量、log P、構造式等）、予測対象物質に割り当てられたQSARクラス、各QSARクラスによる予測結果（予測毒性値、95%予測区間、適用領域の判定）、各QSARクラスの統計値等を得ることができます。

Results

Summary of the Query Chemical

Query Chemical		
SMILES	NCc1ccccc1	
CAS RN [®]	100-46-9	
Chemical Name	phenylmethanamine	
log P	User Input Value <input type="text"/> Re-calculate	
	Estimated Value by KOWWIN™	1.07
	Measured Value in KOWWIN™ Database	1.09
Molecular Weight	107.15	

QSAR Prediction Result

Toxicity filter:

	all	Fish	Daphnid	Alga
Acute	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Chronic	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Statistical filter:

Apply the following statistical criteria:

R² ≥ 0.7 Q² ≥ 0.5 n ≥ 5

Print Detail	QSAR Class Name ¹ <small>Click the class name to see the QSAR details</small>	Type of Predicted Toxicity ²		Predicted Toxicity [mg/L]	95% Prediction Interval	log P (Estimated)	Applicability Domain Judgement		Statistics of QSAR Class					
		Organism	Acute or Chronic				log P ⁴ [Range]	Structure ⁵	R ²	Q ² ⁶	RMSE	n ⁷	criteria ⁸	
<input checked="" type="checkbox"/>	amine primary unreactive NH2=1 aliphatic	Fish	Acute	86	[6.2, 1200]	1.07	in	[-1.61, 5.25]	in	0.84	0.81	0.54	26(2)	✓

Create Print Format

▶ detailed information about the annotations

図5-1 QSAR予測結果画面



(1) 予測対象物質の基本情報

結果画面の上部には入力した SMILES と補足情報が表示されます (図 5-2)。

Query Chemical	
SMILES	NCc1ccccc1 ← b
CAS RN*	100-46-9 ← c
Chemical Name	phenylmethanamine ← d
log P	User Input Value <input type="text"/> Re-calculate ← i
	Estimated Value by KOWWIN™ 1.07 ← f
	Measured Value in KOWWIN™ Database 1.09 ← g
Molecular Weight	107.15 ← h

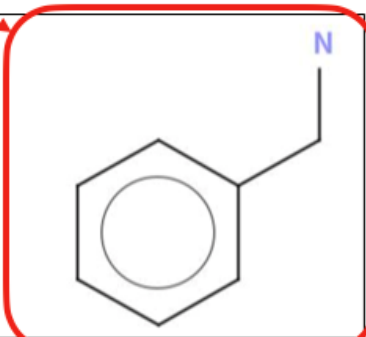


図 5-2 予測対象物質の基本情報表示

- a 予測対象物質の構造式
- b CAS 番号 (ユーザが入力した場合のみ表示)
CAS 番号チェック (チェックディジット*が正しいかどうかの確認のみ) が行われ、チェックディジットが正しくない CAS 番号は、その後ろに「(incorrect)」と表示されます。
※ CAS 番号が正しいかどうかをチェックするための一番右側の桁の数字のこと。入力ミスがないかを検算するのみで、SMILES と CAS 番号が一致しているかを確認するわけではありません。
- c 物質名 (ユーザが入力した場合のみ表示)
- d 予測対象物質の SMILES
- e 予測対象物質の log P のユーザ入力値 (ユーザが入力した場合のみ表示)
- f 予測対象物質の log P の KOWWIN™ による推定値
- g 予測対象物質の log P の実測値
KOWWIN™ データベースにある場合のみ表示されます。複数登録されている場合は、「log P 値(CAS RN: CAS 番号)」の形で、全ての値が表示されます。
- h 予測対象物質の分子量の計算値 (Open Babel プログラムを使用)
- i 「Re-calculate」 ボタン
クリックすると、f 欄に入力されている log P 値を用いて QSAR 予測結果を更新します。



(2) QSAR予測結果

結果画面の中段には、QSARクラス名、予測毒性タイプ、毒性予測結果（緑字）、予測毒性値の95%予測区間、予測対象物質に対して使用したlog P、適用領域の判定、統計値が表示されます（図5-3）。以下で表示内容の説明をします。

Print Detail	QSAR Class Name ¹ Click the class name to see the QSAR details	Type of Predicted Toxicity ²		Predicted Toxicity [mg/L]	95% Prediction Interval	log P (Estimated)	Applicability Domain Judgement		Statistics of QSAR Class					
		Organism	Acute or Chronic				log P ⁴ [Range]	Structure ⁵	R ²	Q ² ⁶	RMSE	n ⁷	criteria ⁸	
<input checked="" type="checkbox"/>	amine primary unreactive NH ₂ =1 aliphatic	Fish	Acute	86	[6.2, 1200]	1.07	in	[-1.61, 5.25]	in	0.84	0.81	0.54	26(2)	✓

図5-3 QSAR 予測結果表示

チェックボックス

a 毒性値フィルタ

チェックされている予測毒性タイプの QSAR クラスのみが表示されます。チェックを外すことにより、表示する予測毒性タイプを絞ることが出来ます。デフォルトでは全てにチェックが入っています。「all」のチェックボックスをチェックすると、対象の右側三つ全てのチェックボックスにチェックが入ります。例えば、「Acute」の「all」チェックボックスにチェックを入れると「Acute」の Fish、Daphnid、及び Alga チェックボックスに同時にチェックが入ります。

b 統計指標フィルタ

統計値 R²、Q²、及び n の条件を満たす QSAR クラスが表示されます。チェックを外したり、値を変更したりすることにより、表示される QSAR クラスを調整することが出来ます。デフォルトでは全てにチェックが入っており、統計指標の基準 (R²≥0.7 and Q²≥0.5 and n≥5) を満たす QSAR クラスが表示されています。

一括印刷フォーマット用

c 全 QSAR クラス一括印刷チェックボックス

全ての QSAR 予測結果を一括印刷する場合にチェックを入れます。初期状態では、全 QSAR クラスが印刷対象 (d の説明参照) となった場合にチェックが入って表示されます。

d 一括印刷フォーマット用チェックボックス

デフォルトでは、QSAR クラスが、統計指標の基準を満たし、適用領域内 (log P 判定が in 且つ、構造判定が in または in(p)) であるものにチェックが入っています。

e 一括印刷フォーマット画面表示用ボタン

クリックすると予測結果画面および d でチェックが入った QSAR クラスの詳細情報を一括印刷するための画面が表示されます (47 ページの「一括印刷フォーマット表示」参照)。



QSAR クラスの名称とリンク

f QSAR クラスの名称

クリックすると、QSAR クラスの詳細情報画面に移動します（28 ページの「QSAR クラス情報の詳細表示」参照）。

予測毒性タイプ

g 生物群 (Fish、Daphnid、もしくは Alga)

h 急性 (Acute) もしくは慢性 (Chronic)

※ 詳細は KATE2025 技術文書参照。

予測値

i 予測毒性値

j 予測毒性値の 95% 予測区間

log P

k 予測対象物質の log P 値

ヘッダの括弧内に、予測対象物質に使用された log P の種類が表示されます。表示文字列の意味は以下の通りとなり、ユーザ入力値が優先して利用されます。

1. User Input : log P のユーザ入力値
2. Estimated : log P の推定値 (KOWWIN™ 利用)

適用領域の判定

l log P 判定結果

この欄の表示内容は、「2. (2)⑤ B) log P 判定」参照。

m [Range]

当該 QSAR クラスのトレーニングセットの記述子 log P の[最小値, 最大値] (log P の適用領域)。

n 構造判定結果

この欄の表示内容は、「2. (2)⑤ A) 構造判定」参照。

統計値

o QSAR 式の R^2

p QSAR 式の Q^2

q QSAR 式の RMSE (平方平均二乗誤差)

r QSAR 式のトレーニングセット数 (サポートケミカルを含まない)

括弧内はサポートケミカルの数。

s 統計指標の基準

基準を満たす場合にチェックマークを表示します。

※ 詳細は KATE2025 技術文書参照。

チェックボックスのチェックを操作する、もしくは統計数値の変更を行うと、即座に表示の更新が行われます。全ての予測毒性タイプに対して QSAR クラスが表示されない場合、「No applicable results.」と表示されます (図 5-4)。



Print Detail <input type="checkbox"/>	QSAR Class Name ^{*1} Click the class name to see the QSAR details	Type of Predicted Toxicity ^{*2}		Predicted Toxicity [mg/L]	95% Prediction Interval	log P (Estimated)	Applicability Domain Judgement		Statistics of QSAR Class				
		Organism	Acute or Chronic				log P ^{*4} [Range]	Structure ^{*5}	R ²	Q ² ^{*6}	RMSE	n ^{*7}	criteria ^{*8}
<div style="border: 2px solid red; padding: 5px; display: inline-block;"> No applicable results. Change the criteria above(R², Q² or n). </div>													

図5-4 QSAR 予測結果表示 (QSAR クラスが1つも表示されない場合)



6. QSARクラス情報の詳細表示 (Verify QSAR画面)

QSAR予測結果でQSARクラス名をクリックすると、そのQSARクラスの詳細情報の画面 (Verify QSAR画面) が表示されます (図6-1~6-3)。そのQSARクラスに関する回帰式のグラフ、トレーニングセットやサポートケミカルの詳細情報、構造クラスの定義、および予測対象物質の部分構造情報等が得られます。

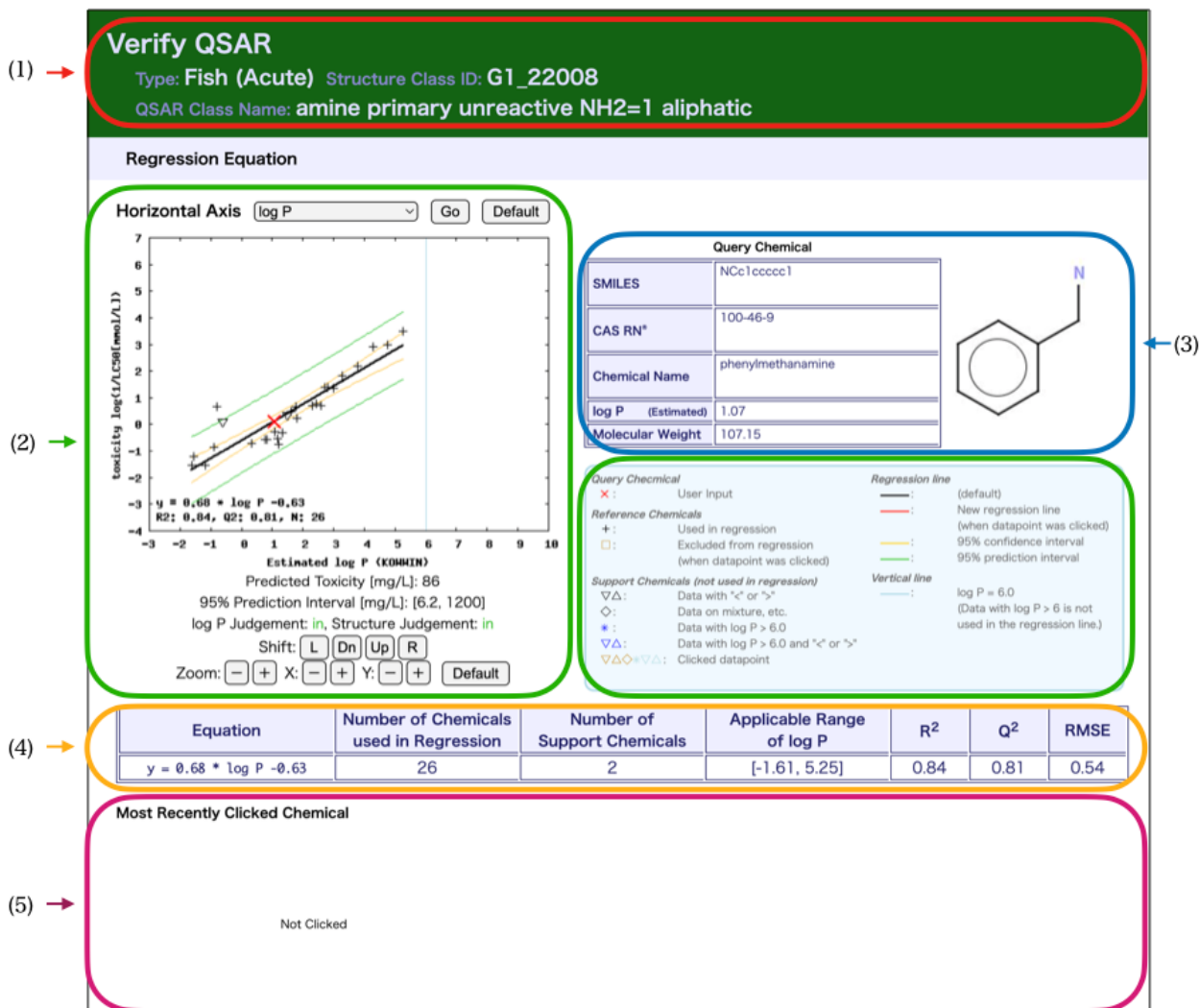


図6-1 Verify QSAR画面 (上段)



(6) → **structure view** | **table view** (8)

List of Chemicals

sort by **X-axis(log P)** in **ascending order** **Update**

Notice:
 - The value below chemical structure represents the similarity (Tanimoto coefficient with PubChem Fingerprints) to the query chemical.
 - The figures in parentheses indicate the coordinate values (x, y) of the chemical in the log P vs. log(1/LC₅₀, EC₅₀, or NOEC) graph.

- Reference chemicals (used in regression)

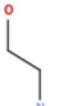
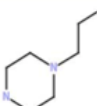
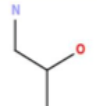
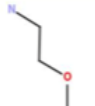
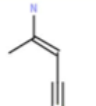
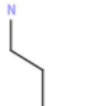
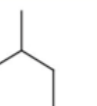
						
0.173 (-1.61, -1.53)	0.211 (-1.57, -1.23)	0.192 (-1.19, -1.53)	0.187 (-0.91, -0.84)	0.240 (-0.79, 0.68)	0.224 (0.34, -0.72)	0.232 (0.76, -0.58)

図 6 - 2 Verify QSAR 画面 (中段)

(9) → **+ Definition of Structure Class**

- Substructures of the Query Chemical

+ Substructures used only for Structural Classification

- Substructures used for the Judgement and the Classification

(10) → Hide SMARTS

Judgement ¹	FragID	Substructure Name	Count	SMARTS
in	5007	Nitrogen [N,n]	1	[#7]
in	5037	pro-SB 1	1	[CH2][NH2]
in	5500	amine (daphnid ACR100)	1	[#7;v3;X3;1S([#7][#6]);1S([#7][#6;X3][#7]);1S([#7][#6]-,#[1#6]);1S([#7][#6;R][1#6;#7;#8;1#16;R][1#6;1#7;#8;1#16;R][1#6;1#7;#8;1#16;R])]

▶ detailed information about the annotation

図 6 - 3 Verify QSAR 画面 (下段)

(1) QSARクラスの基本情報

上部に表示される紺色の帯の各項目は以下のようにになっています (図 6 - 4)。

Verify QSAR

a → **Type: Fish (Acute)** | **Structure Class ID: G1_22008** (b)

c → **QSAR Class Name: amine primary unreactive NH2=1 aliphatic**

図 6 - 4 QSARクラスの基本情報

- a 当該 QSAR クラスに対する予測毒性タイプ
- b 当該 QSAR クラスに対する構造クラス ID
- c QSAR クラスの名称



(2) グラフ関連部分

グラフ関連部分は、以下のようになっています（図6-5）。

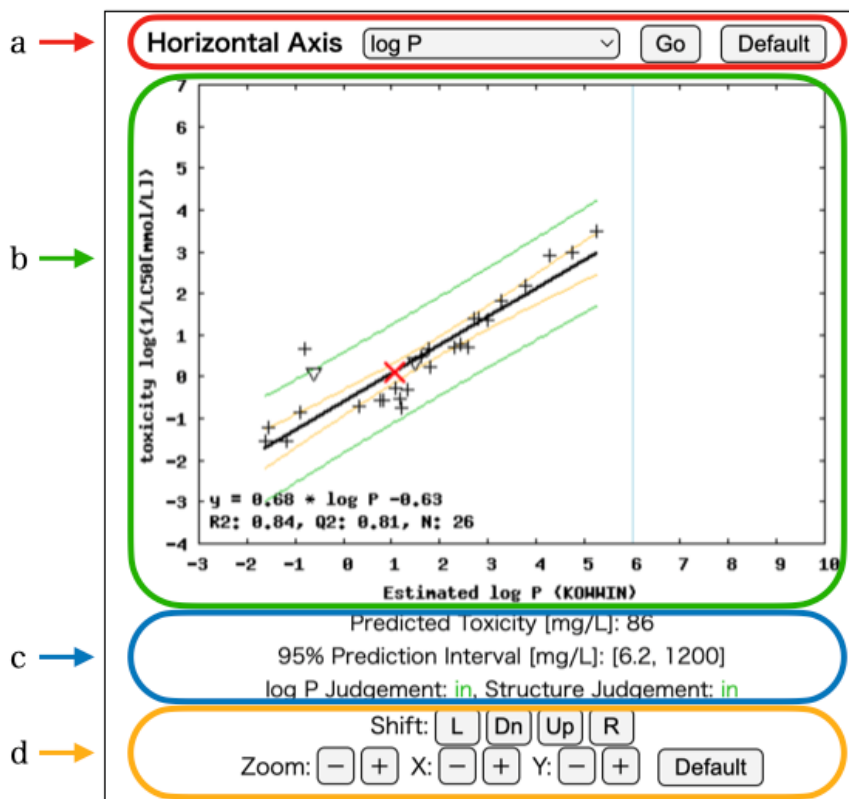


図6-5 グラフ関連（毒性値と log P の関係）

- a 横軸の選択
 デフォルトの「log P」から「Predicted Variable」に変更し、「Go」ボタンをクリックすることにより、 $\log(1/EC_{50}$ 等^{*}[mmol/L])の「実測毒性値 vs. 予測毒性値」のグラフに切り替えることができます。「Default」ボタンをクリックするとデフォルトに戻ります。
^{*} 当該 QSAR クラスの予測毒性のタイプにより、 LC_{50} , EC_{50} , NOEC のいずれかが入ります。本文書では、 EC_{50} 等と記載します。
- b グラフ
 デフォルトでは毒性値 ($\log(1/EC_{50}$ 等 [mmol/L])) と log P の関係をプロットし、左下に QSAR 式と R^2 , Q^2 , 及びトレーニングセットの数を表示します。
- c 予測対象物質情報
 1 行目 予測毒性値 [mg/L]
 2 行目 予測毒性値 [mg/L] の 95% 予測区間
 3 行目 log P 判定と構造判定の結果
- d グラフの移動・縮小・拡大ボタン
 1 行目 Shift L : 左に移動 R : 右に移動
 Dn : 下に移動 Up : 上に移動、
 2 行目 Zoom - : 全体の縮小 + : 全体の拡大
 X - : X 軸のみの縮小 + : X 軸のみの拡大
 Y - : Y 軸のみの縮小 + : Y 軸のみの拡大



グラフ右下に凡例が表示されます (図 6-6)。

Query Chemical		Regression line	
×:	User Input	—:	(default)
Reference Chemicals		—:	New regression line (when datapoint was clicked)
+	Used in regression	—:	95% confidence interval
□:	Excluded from regression (when datapoint was clicked)	—:	95% prediction interval
Support Chemicals (not used in regression)		Vertical line	
▽△:	Data with "<" or ">"	—:	log P = 6.0 (Data with log P > 6 is not used in the regression line.)
◇:	Data on mixture, etc.		
*	Data with log P > 6.0		
▽△:	Data with log P > 6.0 and "<" or ">"		
▽△◇*▽△:	Clicked datapoint		

図 6-6 グラフの凡例

凡例の説明(一部抜粋)

- × : 予測対象物質
- ＋ : トレーニングセットデータ
- : 選択されたトレーニングセットデータ
- 黒色の直線 : 回帰直線
- 赤色の直線 : 「□」を除いて計算した回帰直線
- 橙色の双曲線 : 回帰式の 95%信頼区間
- 緑色の双曲線 : $\log(1/EC_{50}$ 等[mmol/L])の 95%予測区間

物質の選択

「グラフ上のトレーニングセットデータ」(図 6-7 の「+」で表示されるポイント)をクリックして選択すると、選択対象の物質を回帰式の計算から除外した回帰直線(赤色)を追加で表示させることが出来ます。その際、予測区間(緑色の曲線)と信頼区間(橙色の曲線)についても、対象の物質を省いて再計算されます。複数の物質を選択することも可能です。

グラフ上では選択した物質は□で表示されます(図 6-7 の赤枠)。構造式一覧、及び物質データ上では選択された物質は紫の枠に囲まれて表示されます(構造式一覧タブでの例は、図 6-8 の赤枠)。もう一度そのポイントをクリックすると選択が解除されます。

物質の選択を行うと、グラフの左上に、選択した物質数についてのメッセージが表示されます(図 6-7 の青枠)。また、グラフ内の右下(図 6-7 の緑枠右側)には「→」印の後に物質を削除後の QSAR 式と R^2 、 Q^2 、及びトレーニングセットの数が表示され、左側のデフォルトでの値と比較する事が可能です。



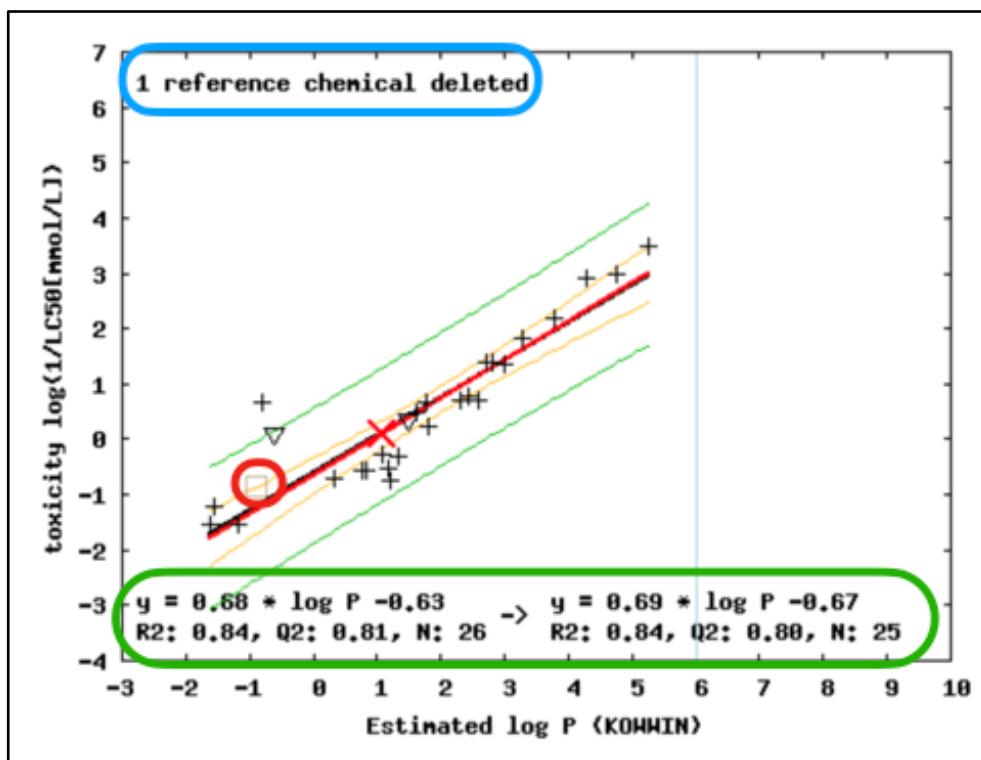


図 6-7 物質の選択 (グラフ)

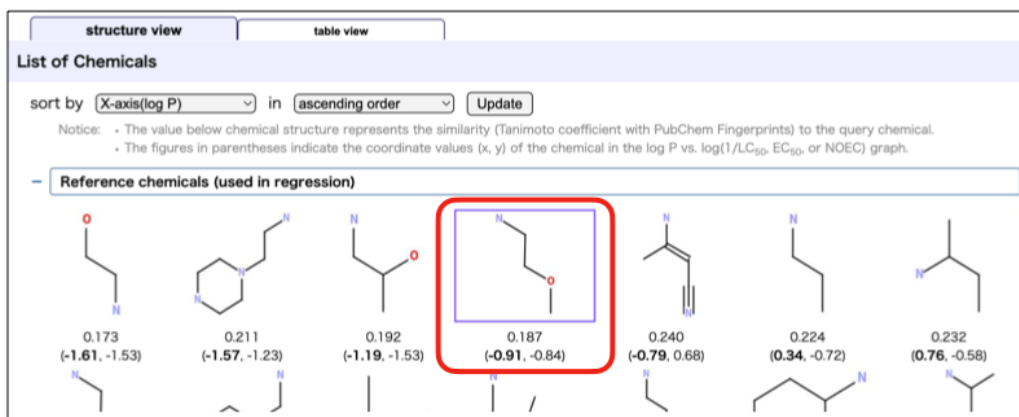


図 6-8 物質の選択 (構造式一覧での例)

構造式一覧、及び物質データに表示されている構造式 (図 6-8 の構造式、後述の図 6-15 の中央等) についてはクリック可能で、上述のグラフを含めた三者で連動します。例えば、図 6-8 の任意の構造式をクリックすれば、グラフ上、及び物質データの選択状態は相応に遷移します。

また、後述する構造式一覧でのソート機能でソートしても、選択状態は保持されたままソートされます。



(3) 予測対象物質情報

グラフの右側上部に、予測対象物質の基本的な情報、及び構造式が表示されます（図 6-9）。

Query Chemical	
SMILES	<chem>NCc1ccccc1</chem>
CAS RN*	100-46-9
Chemical Name	phenylmethanamine
log P (Estimated)	1.07
Molecular Weight	107.15

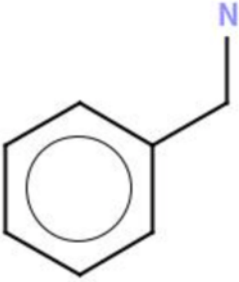


図 6-9 予測対象物質情報

表内の各行は以下のようになっています。

SMILES : 予測対象物質の SMILES

CAS RN : 予測対象物質に対するユーザが入力した CAS 番号

Chemical Name : 予測対象物質に対するユーザが入力した物質名

log P : 予測対象物質の log P 値。値の後ろには、ユーザ入力値の場合「(User Input)」、KOWWIN による推定値の場合「(Estimated)」が表示される。

Molecular Weight : 予測対象物質の分子量

(4) 回帰式情報

グラフ関連部分の下に回帰式 (QSAR式) 情報が表示されます (図 6-10)。

Equation	Number of Chemicals used in Regression	Number of Support Chemicals	Applicable Range of log P	R ²	Q ²	RMSE
$y = 0.68 * \log P - 0.63$	26	2	[-1.61, 5.25]	0.84	0.81	0.54

図 6-10 回帰式情報

各列は以下のようになっています。

Equation : 回帰式 (QSAR 式)

Number of Chemicals used in Regression : トレーニングセット数

Number of Support Chemicals : サポートケミカル数

Applicable Range of log P : トレーニングセット中の log P 範囲 (最小値と最大値)

R² : QSAR 式の R²

Q² : QSAR 式の Q²

RMSE : QSAR 式の RMSE



(5) 最後にクリックされた物質

「Most Recently Clicked Chemical」の下に「最後にクリックされた物質」の情報が表示されます（図6-11、図6-12）。初期状態では「Not Clicked」と表示されています（図6-11）。

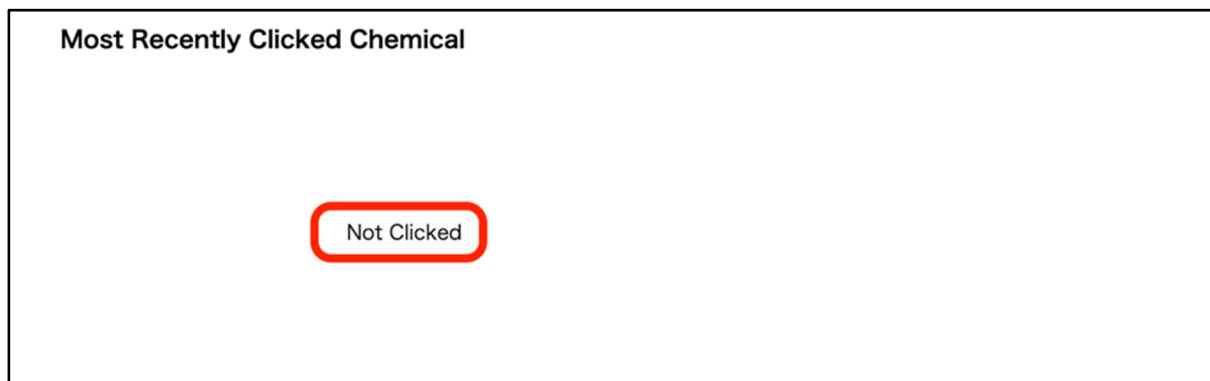
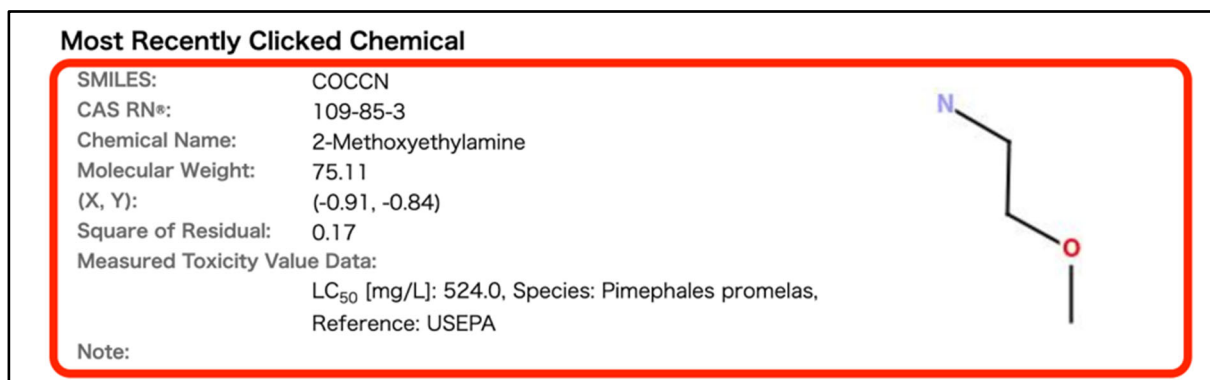


図6-11 最後にクリックした物質の情報（物質のクリック前）

物質をクリックすると「最後にクリックされた物質」に関する情報と構造式が表示されます（図6-12）。

A screenshot of a software interface showing detailed information for a chemical. The title 'Most Recently Clicked Chemical' is at the top. Below it, a red-bordered box contains the following text:

SMILES:	COCCN
CAS RN#:	109-85-3
Chemical Name:	2-Methoxyethylamine
Molecular Weight:	75.11
(X, Y):	(-0.91, -0.84)
Square of Residual:	0.17
Measured Toxicity Value Data:	LC ₅₀ [mg/L]: 524.0, Species: Pimephales promelas, Reference: USEPA
Note:	

To the right of the text is a chemical structure diagram of 2-methoxyethylamine, showing a nitrogen atom (N) at the top, a two-carbon chain, and a methoxy group (-OCH₃) at the end.

図6-12 最後にクリックした物質の情報（物質のクリック後）

各行は以下のようにになっています。

- SMILES
- CAS 番号
- 物質名
- 分子量
- グラフ上の座標
- 残差の二乗
- 実測毒性値、生物種、出典
- 補足情報（試験に関する情報等がある場合のみ）



(6) QSARクラスに含まれる物質情報の表示切り替え機構

Verify QSAR画面中段では、該当QSARクラスのトレーニングセット及びサポートケミカルの情報を、構造式一覧 (structure viewタブ) と物質データ (table viewタブ) に分けて記載しています。タブをクリックすることにより表示方法を切り替えることができます。

(7) 構造式一覧 (structure viewタブ)

structure viewタブ (図6-13の赤枠アクティブ時) には、このQSARクラスのトレーニングセットおよびサポートケミカルの構造式一覧が表示されます (図6-13)。

structure view | table view

List of Chemicals

sort by X-axis(log P) in ascending order Update

Notice: The value below chemical structure represents the similarity (Tanimoto coefficient with PubChem Fingerprints) to the query chemical.
The figures in parentheses indicate the coordinate values (x, y) of the chemical in the log P vs. log(1/LC₅₀, EC₅₀, or NOEC) graph.

- Reference chemicals (used in regression)

0.173	0.211	0.192	0.187	0.240	0.224	0.232
(-1.61, -1.53)	(-1.57, -1.23)	(-1.19, -1.53)	(-0.91, -0.84)	(-0.79, 0.68)	(0.34, -0.72)	(0.76, -0.58)
0.254	1.000	0.257	0.254	0.284	0.324	0.873
(0.83, -0.56)	(1.07, -0.29)	(1.18, -0.51)	(1.21, -0.74)	(1.33, -0.31)	(1.63, 0.48)	(1.76, 0.67)
0.299	0.309	0.304	0.247	0.292	0.300	0.851
(1.82, 0.25)	(2.31, 0.72)	(2.43, 0.78)	(2.58, 0.72)	(2.73, 1.40)	(2.80, 1.40)	(3.00, 1.34)
0.300	0.300	0.300	0.300	0.300		
(3.29, 1.82)	(3.78, 2.18)	(4.27, 2.91)	(4.76, 2.98)	(5.25, 3.49)		

+ Support Chemicals (not used in regression)

図6-13 物質データの構造式一覧

a ソート機能

左のプルダウンでソート項目を選択、右のプルダウンで昇順か降順を選択し、「Update」をクリックすることでソートすることが出来ます。ソート対象は以下となります。

1. 予測対象物質と当該物質の類似度 (Similarity)
PubChem fingerprint を用いた Tanimoto 係数 (詳細は KATE2025 技術文書参照)。数値は0から1の範囲内で、予測対象物質との類似性が高いものほど1に近い値を取る。
2. 当該物質のグラフ上の X 座標 (X-axis(log P))
3. 当該物質のグラフ上の Y 座標 (Y-axis(toxicity))



$\log(1/EC_{50}$ 等 [mmol/L])の値。

4. 当該物質の残差の二乗 (Square of Residual)

類似度、X座標及びY座標でソートした場合には、ソートされた該当項目(図6-12のc, d内)が、太字で強調表示されます。デフォルトではX座標 ($\log P$) の昇順でソートされています。

b Reference chemicals (used in regression)

トレーニングセットの構造式一覧。

c 予測対象物質との類似度

d グラフ上の座標 (X座標、Y座標)。

e Support Chemicals (not used in regression)

サポートケミカルの構造式一覧 (デフォルトでは折り畳まれています)。

※ KATE2025 では、タイトル左横に「+」「-」の記載がある項目は、展開・折り畳みが可能です (図6-13のbやe、後述図6-15など)。

「Support Chemicals (not used in regression)」 (図6-13のe) をクリックすると、サポートケミカルの構造式一覧が、物質の情報に基づき、分けられた形で表示されます (図6-14)。各構造式に対する情報は、前出のトレーニングセットと同じ形式です (図6-13の説明参照)。

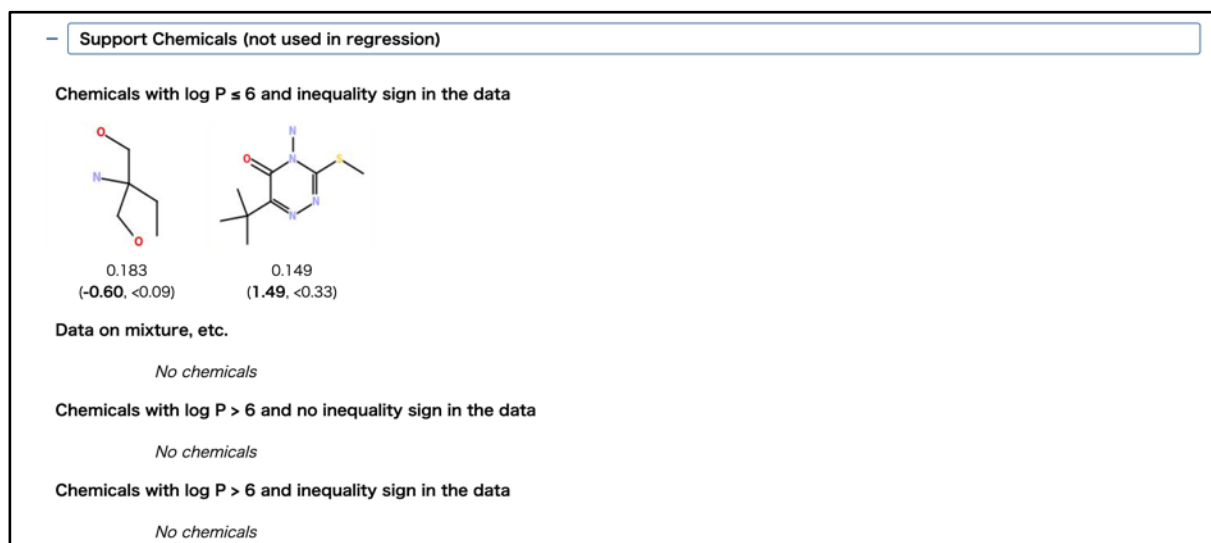


図6-14 サポートケミカルの構造式一覧



(8) 物質データ(table viewタブ)

table viewタブ（図6-15の赤枠アクティブ時）には、このQSARクラスのトレーニングセットおよびサポートケミカルの詳細情報が表形式で表示されます（図6-15）。

内容は構造式一覧と同様に、「Reference chemicals (used in regression)」の下に当該QSARクラスに属するトレーニングセットの詳細情報一覧、「Support Chemicals (not used in regression)」の下に当該クラスに属するサポートケミカルの詳細情報一覧が表示されます。

CAS RN*	Chemical Name	SMILES	Structural Formula	Similarity	Molecular Weight	Estimated log P	Measured Toxicity Data			
							LC ₅₀ [mg/L]	log(1/LC ₅₀) [mmol/L]	Reference	Note
141-43-5	Monoethanolamine	NCCO		0.173	61.08	-1.61	2070.0	-1.53	USEPA	
140-31-8	1-(2-Aminoethyl)piperazine	NCCN1CCNCC1		0.211	129.21	-1.57	2190.0	-1.23	USEPA	
78-96-6	1-Amino-2-propanol	CC(O)CN		0.192	75.11	-1.19	2520.0	-1.53	USEPA	

図6-15 物質データ

- a CAS RN : CAS 番号
- b Chemical Name : 物質名
- c SMILES : KATE2025 で使用されている SMILES
- d Structural Formula : 構造式
- e Similarity : 当該物質と予測対象物質との類似度
- f Molecular Weight : KATE2025 で使用されている分子量
- g Estimated log P : KOWWIN™による log P 推定値
- h LC₅₀*[mg/L] : 毒性値 (EC₅₀ 等[mg/L])
- i log(1/LC₅₀*[mmol/L]) : 毒性値 (log(1/EC₅₀ 等[mmol/L]))
 ※ h と i については、当該 QSAR クラスの予測毒性のタイプにより、LC₅₀, EC₅₀, NOEC のいずれかが表示されるようになっています。
- j Reference : 元データの出典（「MOE 試験実施年度」（例: MOE 2007）または「USEPA」）。ここで「MOE」のデータの出典は以下の URL となっています。
<https://www.env.go.jp/content/000212330.pdf>
 （環境省 化学物質の環境リスク評価 第22巻 第2編 「(II-1) 環境省の生態影響試験（藻類、甲殻類、魚類）結果一覧（令和6年3月版）」。2025年3月1日アクセス）



「USEPA」のデータの出典は、米国環境保護庁（US EPA）のファットヘッドミノール・データベースの魚類急性毒性試験結果となっています（下記URL参照。2025年3月1日アクセス）。

https://archive.epa.gov/med/med_archive_03/web/html/fathead_minnow.html

k Note：物質に関するその他の情報（生態影響試験に関する情報等）

表のソート機能

KATE2025の一部の表には、テーブルヘッダをクリックする事によるソート機能を実装しています。ソート可能な列には、ヘッダ右下に、上下の三角マークが表示されています（図6-16の赤枠）。ヘッダをクリックする毎に昇順降順が切り替わり、対応した三角が塗り潰されます（図6-16の場合は「Estimated log P」列）。


Structural Formula	Similarity	Molecular Weight	Estimated log P	LC ₅₀	log(
				[mg/L]	[mm
	0.172	61.09	1.61	2070.0	

図6-16 表のソート機能

- ※ Shift キーを押しながらクリックすることで、複数列でのソートが可能です。
- ※ 一部の表では、ソート出来ない列のヘッダをクリックすることで初期状態に戻す事が可能です。

(9) 構造クラス定義

「Definition of Structure Class」をクリックすると、当該QSARクラスに対応する構造クラスの定義情報が表示されます（図6-17）。

割り当てられた構造クラスが黄色でハイライトされており、「Structure ID」列に構造クラスID、「Description」列に構造クラスもしくは部分構造の名称、「Decision Tree」列に構造クラスもしくは部分構造の定義が記載されています。

Definition of Structure Class		
Group: Amine primary Show all structures		
Structure ID	Description	Decision tree
-	amine CNH2	ID:3100 > 0
-	amine primary unreactive	L R_00033 = false
-	NH2 amine unreactive	L G1_00010 = false
-	amine primary unreactive NH2=1	L ID:3100 = 1
G1_22008	amine primary unreactive NH2=1 aliphatic	L ID:4510 = 0
GA_22008	amine primary unreactive NH2=1 aliphatic alga	L RA_00033 = false

図6-17 構造クラス定義情報



(10) 予測対象物質の部分構造一覧

「Substructures of the Query Chemical」の下の「Substructures used for the Judgement and the Classification」の下には「構造分類と構造判定の両方に使用される部分構造一覧」が表示されます（図6-18）。

f →

Judgement ¹	FragID	Substructure Name	Count	SMARTS
in	5007	Nitrogen [N,n]	1	[#7]
in	5037	pro-SB 1	1	[CH2][NH2]
in	5500	amin (daphnid ACR100)	1	[#7;v3;X3;!S([#7][!#6]);!S([#7][!#6;X3]([#7])([#7]);!S([#7][!#6]-#(!#6));!S([#7][!#6;R][!#6;!#7;!#8;!#16;R][!#6;!#7;!#8;!#16;R][!#6;!#7;!#8;!#16;R])]

a b c d e

図6-18 予測対象物質の部分構造一覧（構造分類と構造判定の両方に使用）

- a Judgement：構造に対する適用領域の判定結果（KATE2020 version 2.0からの新機能）の詳細情報であり、QSARクラスの構造判定が「out」や「in(p)」であった場合に、予測対象物質に含まれる部分構造のうちどれがその判定結果の原因であるかがわかります。判定には以下があります。

in：適用領域内

当該部分構造が、当該 QSAR クラスの「構造判定用部分構造リスト」に含まれる場合。

in(p)：条件付き適用領域内

「in」の条件には合致しないが、当該部分構造が、当該 QSAR クラスの「構造判定用部分構造リスト」、あるいは Narcotic Group クラスの「構造判定用部分構造リスト」に含まれる場合。

out：適用領域外

「in」と「in(p)」の何れの条件にも合致しない場合。すなわち当該部分構造に、当該 QSAR クラスの「構造判定用部分構造リスト」と Narcotic Group クラスの「構造判定用部分構造リスト」の何れにも含まれない部分構造がある場合。

- b FragID：部分構造 ID。KATE 用に設定した4桁の番号であり、番号は開発の便宜上、任意に設定したものです。先頭が5で始まります（詳細は KATE2025 技術文書参照）。
- c Substructure Name: 部分構造の名称。
- d Count：予測対象物質に対する SMARTS 合致数（当該部分構造の個数）。
- e SMARTS：部分構造の定義。
- f SMARTS の表示／非表示の切り替えボタン：デフォルトでは「Hide SMARTS」で、クリックすると「SMART」列が非表示となり、ボタン名が「Display SMARTS」に変わります。もう一度クリックすると、再び SMARTS 列が表示され、ボタン名が「Hide SMARTS」に切り替わります。

また、「Substructures used only for Structural Classification」をクリックすると、「構造分類のみに使用される部分構造一覧」が表示されます（図6-19）。



– Substructures of the Query Chemical

– Substructures used only for Structural Classification

FragID	Substructure Name	Count	SMARTS
3001	elements other than CX	1	[!#6;!#9;!#17;!#35;!#53]
3003	elements other than COX	1	[!#6;!#8;!#9;!#17;!#35;!#53]
3004	elements other than CSX	1	[!#6;!#16;!#9;!#17;!#35;!#53]
3009	elements other than COSX	1	[!#6;!#8;!#16;!#9;!#17;!#35;!#53]
3011	elements other than COns	1	[!#6;!F;!Cl;!Br;!I;!n;!s;!o;!O]
3014	elements other than CnosX	1	[\$([!#6;!F;!Cl;!Br;!I;!n;!s;!o]),\${(n+)}]
3022	Carbon	7	[#6]
3100	amine CNH2	1	[#7X3H2;!\${(#7)[*v6]};!\$(N[#6](-[#7,#8,#16]))]
3121	amine Nv3 not hindered	1	[#7v3X3;!\${(NR0)[CR1][CR1]({CX4R0})[CX4R1]};!\${(NR1)(C)C(C)(C)C)};!\${(#7)[#7]};!\$(NC(=[CH2]))};!\$(N[#6](-[#7,#8,#16]))]
4543	MF: not C,c,O,F	1	[!C;!c;!O;!F]
4711	aliphatic-NH2	1	[N;H2;v3;X3;!\$(NC=[S,N,O]);!\$(NCC(=O)O)] [C]
4892	MF: not CHO (kPilotO)	1	[!C;!c;!O]
4893	MF: not CHOP	1	[!C;!c;!O;!P]
4910	aromatic	6	[a]

a

b

c

d

図 6 – 1 9 予測対象物質の部分構造一覧（構造分類のみに使用）

- a FragID：部分構造 ID。KATE 用に設定した 4 桁の番号であり、番号は開発の便宜上、任意に設定したものです。先頭が 3,4,6,7 のものが存在します（詳細は KATE2025 技術文書参照）。
- b Substructure Name：部分構造の名称。
- c Count：予測対象物質に含まれる当該部分構造の個数）。
- d SMARTS：部分構造の定義。



7. 複数の化学物質の予測 (Upload 画面)

KATE2025では、後述のファイルフォーマットに準じて記載された、複数のSMILESを含むファイルをアップロードすることにより、複数の化学物質に対するQSAR予測を一度に行うことが出来ます。一度に予測できる物質数の上限は、2025年現在、1000物質となっています。

Front Page中段の「Prediction of Multiple Chemicals」項からご利用下さい。

(1) 入力ファイルフォーマットについて

KATE2025では、複数のSMILESを連続して処理するために、各行に1物質のデータの連続からなる入力ファイルフォーマット「SMILES List」を使用します。ファイル名自体に制限はありません。

SMILES Listフォーマットは、以下の様に定義されています。

1. タブ区切りテキスト形式です。
2. 1行目に列名の記載が必要ですが、「SMILES (SMILES)」、「CAS (CAS番号)」、「NAME (物質名)」、「LOGP (入力log P値)」、そして「ID(ユーザ定義のID)」以外の文字列は受け付けません。大文字小文字は任意です。
3. 全ての行にはSMILESの記載が必要で、他の列は任意入力です。
4. ID列は、半角英数字と「-」(半角ハイフン)と、「_」(半角アンダースコア)のみ使用できます。
5. 全ての行でタブの数が同じである必要があります。
6. 列は順不同です。

以下に、2つ例を記載します。

例1 : SMILES 列のみ

```
SMILES
CCCCOC (=O) CS
CC (=C) CS
CC1 (CC2 (C) CC3 (Br) C1) CC (Br) (C2) C3
CCCCCCCCBr
CCCCCCCC (Br) CBr
CCCCCCCCBr
```

例2 : 任意の列を含む

ID	NAME	LOGP	SMILES	CAS
A01	name1	-0.8	CCCCOC (=O)	
A20	name2		CC (=C) CS	
B06	name3	1.3	NCC1CCNC1	3731-52-0

※ 「■」と、「■」はタブを示しています。

IDがA01のCAS番号は未入力、A20のlog PとCAS番号も未入力です。

SMILES Listフォーマットについての説明は、「Prediction of Multiple Chemicals」下部の「about the SMILES List format」をクリックすることでも、確認することができます (図7-1)。



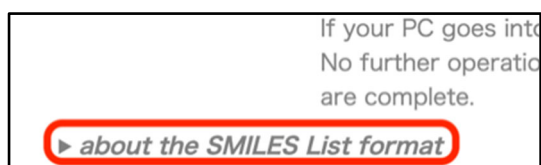


図 7-1 SMILES List フォーマット説明記載場所

(2) 予測実行手順

Front Page 中段の「Prediction of Multiple Chemicals」項の「Select」ボタンをクリックします (図 7-2)。

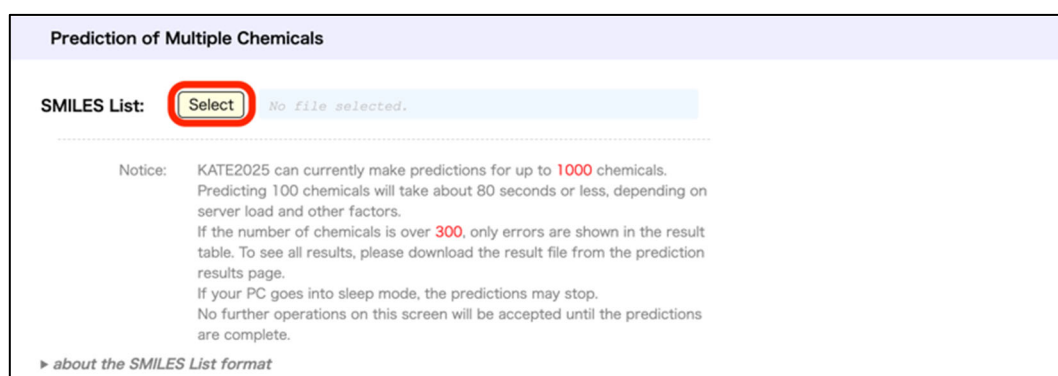


図 7-2 「複数の化学物質予測」用の Select ボタン

毒性予測を行いたい化学物質のSMILESが書かれているファイルを選択して「開く」をクリックします (図 7-3)。

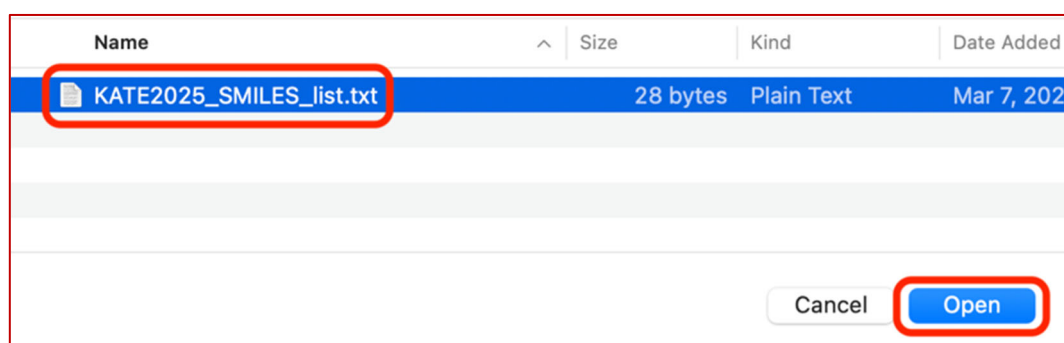


図 7-3 SMILES list の選択

ここで、ファイルに以下の様な不備がある場合には、アラートが出ます。

1. テキストファイル以外が与えられた場合
2. ファイルの記載がSMILES listフォーマットに合わない場合 (列名が間違っている、各行でタブの数が一致しない等)
3. 予測上限 (1,000) を越えた物質数が記載されている場合

ファイルがアップロードされると、「Select」ボタンの右側にファイル名と物質数が表示され、更なる右に「Predict」ボタンが表示されます。このボタンをクリックすると予測が開始します (図 7-4)。



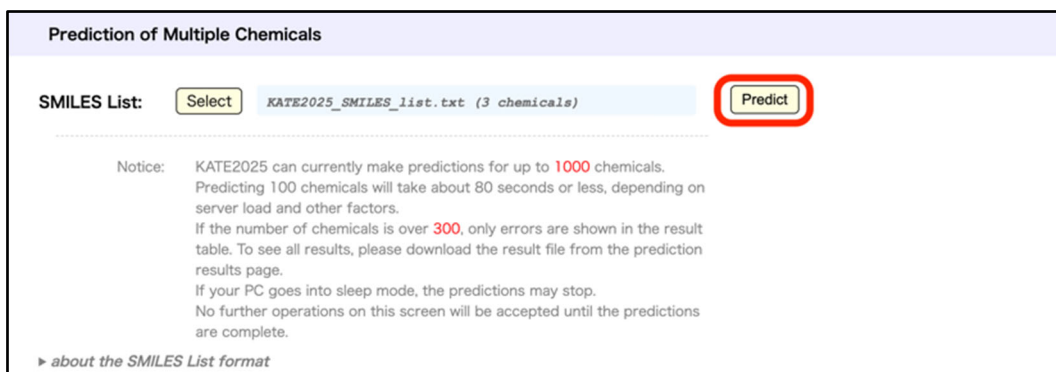


図 7 - 4 「複数の化学物質予測」用の予測 (Predict) ボタン

予測が開始すると画面右に進捗を示すバー (プログレスバー) が表示されます。(図 7 - 5)。プログレスバーの上には「Total time」が表示され、予測の完了までに必要なおおよその時間が表示されます (予測の進捗とともにカウントダウンはしません)。この時間は与えられた物質数から算出したもので、正確な計算時間を示すものではありません。予測時間は化学物質の構造によって変化しますが、通常は100物質の計算であれば約80秒以内に終了します。

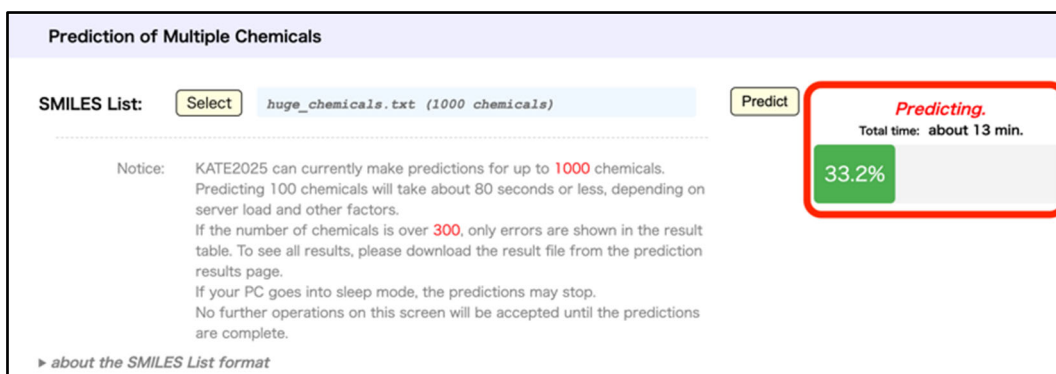


図 7 - 5 計算の進捗状



(3) QSAR予測結果（複数の化学物質）

計算終了後、予測結果が表示されます（図7-6）。

Results (batch mode)

about the result file associated with the button in the right. Download results as text file

QSAR Prediction Results: 0 errors / 3 chemicals

Toxicity filter: all Fish Daphnid Alga

Statistical filter: Apply the following statistical criteria:
 R² ≥ 0.7 Q² ≥ 0.5 n ≥ 5

No	ID	CAS RN*	Chemical Name	SMILES	Molecular Weight	Structural Formula	QSAR Class Name*1 <small>Click the class name to see the QSAR details</small>	Type of Predicted Toxicity*2		Predicted Toxicity [mg/L]	95% Prediction Interval	log P (Estimated)	Applicability Domain Judgement		Statistics of QSAR Class												
								Organism	Acute or Chronic				log P*4 [Range]	Structure*5	R ²	RMSE	n ⁷	criteria*8									
1	1			CCC	44.10		C_X hydrocarbon unreactive aliphatic w/o X	Fish	Acute	36	[4.7, 280]	1.81	out	[2.58, 4.98]	in	0.73	0.68	0.36	21(4)	✓							
							narcotic group Fish Acute	Fish	Acute	66	[9.1, 480]	1.81	in	[-0.63, 5.88]	in	0.87	0.87	0.43	151(34)	✓							
							narcotic group Daphnid Acute	Daphnid	Acute	11	[1.5, 87]	1.81	in	[1.08, 5.88]	in	0.71	0.70	0.43	81(24)	✓							
							narcotic group Alga Acute	Alga	Acute	61	[7.2, 510]	1.81	in	[1.08, 5.26]	in	0.75	0.73	0.44	50(47)	✓							
							Cnos_X unreactive Fish Chronic	Fish	Chronic	0.99	[0.073, 13]	1.81	in	[1.52, 5.52]	in	0.76	0.68	0.43	12(0)	✓							
							C_X hydrocarbon unreactive	Fish	Chronic	1.1	[0.077, 14]	1.81	in	[1.52, 5.52]	in	0.78	0.68	0.43	11(0)	✓							
							narcotic group Fish Chronic	Fish	Chronic	1.1	[0.093, 13]	1.81	in	[1.52, 5.81]	in	0.82	0.75	0.41	12(0)	✓							
							narcotic group Daphnid Chronic	Daphnid	Chronic	0.99	[0.082, 12]	1.81	in	[-1.20, 5.88]	in	0.70	0.68	0.54	73(14)	✓							
							2	2			CCN	45.08		amine primary unreactive NH2=1 aliphatic	Fish	Acute	240	[17, 3500]	-0.15	in	[-1.61, 5.25]	in	0.84	0.81	0.54	26(2)	✓
														narcotic group Fish Acute	Fish	Acute	3600	[480, 27000]	-0.14	in	[-0.63, 5.88]	in	0.87	0.87	0.43	151(34)	✓

図7-6 「複数の化学物質予測」の予測結果

チェックボックスと表の内容は1つの化学物質予測の場合とほぼ同じ項目となっています。QSAR Class Name列のQSARクラス名をクリックすると、Verify QSAR画面に移動します。

複数化学物質の予測では、計算が出来ないSMILESがあった場合は計算をスキップし、エラーとして集計します。「QSAR Prediction Results」と表示されたタイトル内右横には、そのエラー件数、及び全件数が表示されます（図7-7の a）。エラーがあった場合には、エラー件数部分は赤字となります（図7-8）。また、結果の表中の該当行には、エラーメッセージを表示します。エラーが発生したもののみの情報を確認したい場合は、cをクリックし表示させて下さい（図7-7の b, c、図7-8）。

QSAR Prediction Results: **0 errors / 3 chemicals**

Toxicity filter: all Fish Daphnid Alga

Statistical filter: Apply the following statistical criteria:
 R² ≥ 0.7 Q² ≥ 0.5 n ≥ 5

all chemicals | chemicals with errors only

Table header: QSAR Class Name*1 | Type

図7-7 「複数の化学物質予測」結果タブ機構

- a エラー数／総物質数表示
 エラーがあった場合は、エラー数が赤字で表示される。



- b 「全物質」タブ
 デフォルトではこちらが表示される。毒性値フィルタ、統計指標フィルタによる絞り込み表示が可能。
- c 「エラーのみ」タブ
 エラーのみを閲覧したい場合に使用（図7-8）。フィルタ機構はグレーとなり、絞り込みの操作は出来ない。

Results (batch mode)

about the result file associated with the button in the right. [Download results as text file](#)

QSAR Prediction Results: **2 errors / 5 chemicals**

Toxicity filter: all Fish Daphnid Alga
 Acute
 Chronic

Statistical filter: Apply the following statistical criteria:
 $R^2 \geq 0.7$ $Q^2 \geq 0.5$ $n \geq 5$

No	ID	CAS RN*	Chemical Name	SMILES	Molecular Weight	Structural Formula	QSAR Class Name ¹ <small>Click the class name to see the QSAR details</small>	Type of Predicted Toxicity ² Organism Acute or Chronic	Predicted Toxicity [mg/L]	95% Prediction Interval	log P ³	Applicability Domain Judgement		Statistics of QSAR Class				
												log P ⁴ [Range]	Structure ⁵	R ²	Q ² * ⁶	RMSE	n ⁷	criteria ⁸
2	test1		test	NCc1%12>^ccnc1		Image Not Available	[SMILES] '1' are prohibited on KATE2025. (count: 1) [SMILES] '**' are prohibited on KATE2025. (count: 1) [SMILES] '.,' are prohibited on KATE2025. (count: 1) [SMILES] '>' are prohibited on KATE2025. (count: 1)				Error!!							
4	test2		test	NCc1ccnc1A		Image Not Available	[SMILES] 'A' : 'A' (except '[As]') should not be included. [log P] Invalid numerical value for log P: STR				Error!!							

Please download and check the result file from the button in the top right.

detailed information about the annotations

図7-8 エラーが存在した場合の「エラーのみ」タブ表示例

化学物質が300物質より多い場合は、エラーのあった物質のみの情報が表示され、エラーのない物質も含めた全物質の予測結果を表示することはできません（図7-9）。

Results (batch mode)

about the result file associated with the button in the right. [Download results as text file](#)

QSAR Prediction Results: **0 errors / 1000 chemicals**

Toxicity filter: all Fish Daphnid Alga
 Acute
 Chronic

Statistical filter: Apply the following statistical criteria:
 $R^2 \geq 0.7$ $Q^2 \geq 0.5$ $n \geq 5$

Because the number of chemicals is over 300, only errors are shown in the result table.

No	ID	CAS RN*	Chemical Name	SMILES	Molecular Weight	Structural Formula	QSAR Class Name ¹ <small>Click the class name to see the QSAR details</small>	Type of Predicted Toxicity ² Organism Acute or Chronic	Predicted Toxicity [mg/L]	95% Prediction Interval	log P (Estimated)	Applicability Domain Judgement		Statistics of QSAR Class			
												log P ⁴ [Range]	Structure ⁵	R ²	Q ² * ⁶	RMSE	n ⁷
No errors found.																	

Please download and check the result file from the button in the top right.

detailed information about the annotations

図7-9 予測物質が300より多い場合

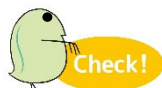
化学物質の数が上述閾値より多い場合や、結果一覧を手元に保存したい場合は、画面右上の「Download results as tsv」をクリックすることにより、予測結果一覧をタブ区切りテキスト形式のファイルでダウンロードすることができます。このファイルには全ての物質の予測結果が含まれます（図7-10）。



ID	CAS RN	Chemical Name	SMILES	Molecular Weight	QSAR ID	QSAR Class Name	Organism	Acute or Chronic	Predicted Toxicity	95% Prediction Interval	log P Value	log P Type	log P Judgement	log P Range	Structure Judgement	R2	Q2	RMSE	n	criteria
1			COc	44.1	12120341	C,X hydrocarbon unreactive aliphatic w/o X	Fish	Acute	36 [4.7, 280]		1.81	Estimated	out	[2.58, 4.98] in		0.73	0.68	0.36	21(4)	yes
1			COc	44.1	12899941	narcotic group Fish Acute	Fish	Acute	66 [9.1, 480]		1.81	Estimated	in	[-0.63, 5.88] in		0.87	0.87	0.43	151(34)	yes
1			COc	44.1	22120341	C,X hydrocarbon unreactive aliphatic w/o X	Daphnid	Acute	9.3 [0.67, 130]		1.81	Estimated	out	[2.58, 5.74] in		0.66	0.51	0.42	15(3)	no
1			COc	44.1	22899941	narcotic group Daphnid Acute	Daphnid	Acute	11 [1.5, 87]		1.81	Estimated	in	[1.08, 5.88] in		0.71	0.7	0.43	81(24)	yes
1			COc	44.1	32120341	C,X hydrocarbon unreactive aliphatic w/o X	Alga	Acute	370 [1.6, 87000]		1.81	Estimated	out	[2.58, 4.08] in		0.66	0.38	0.47	6(10)	no
1			COc	44.1	32899941	narcotic group Alga Acute	Alga	Acute	61 [7.2, 510]		1.81	Estimated	in	[1.08, 5.26] in		0.75	0.73	0.44	50(47)	yes
1			COc	44.1	12100151	C,X hydrocarbon unreactive	Fish	Chronic	1.1 [0.077, 14]		1.81	Estimated	in	[1.52, 5.52] in		0.78	0.68	0.43	11(0)	yes
1			COc	44.1	12500151	Cros,X unreactive Fish Chronic	Fish	Chronic	0.99 [0.073, 13]		1.81	Estimated	in	[1.52, 5.52] in		0.76	0.68	0.43	12(0)	yes
1			COc	44.1	12899851	narcotic group Fish Chronic	Fish	Chronic	1.1 [0.093, 13]		1.81	Estimated	in	[1.52, 5.81] in		0.82	0.75	0.41	12(0)	yes
1			COc	44.1	22120351	C,X hydrocarbon unreactive aliphatic w/o X	Daphnid	Chronic	1.3 [0.067, 24]		1.81	Estimated	out	[2.58, 5.74] in		0.58	0.32	0.44	11(2)	no
1			COc	44.1	22899851	narcotic group Daphnid Chronic	Daphnid	Chronic	0.99 [0.082, 12]		1.81	Estimated	in	[-1.20, 5.88] in		0.7	0.68	0.54	73(14)	yes
1			COc	44.1	32100151	C,X hydrocarbon Alga Chronic	Alga	Chronic	2.8 [0.14, 54]		1.81	Estimated	in	[1.61, 5.52] in		0.38	0.32	0.61	55(25)	no
1			COc	44.1	32120351	C,X hydrocarbon unreactive aliphatic w/o X	Alga	Chronic	82 [0.26, 26000]		1.81	Estimated	out	[2.58, 4.20] in		0.51	0.24	0.62	9(9)	no
1			COc	44.1	32899851	narcotic group Alga Chronic	Alga	Chronic	6.2 [0.23, 160]		1.81	Estimated	in	[0.69, 5.81] in		0.57	0.53	0.69	57(27)	no
2			COc	45.08	12200841	amine primary unreactive NH2-1 aliphatic	Fish	Acute	240 [17, 3500]		-0.15	Estimated	in	[-1.61, 5.25] in		0.84	0.81	0.54	26(2)	yes
2			COc	45.08	22200841	amine primary unreactive NH2-1 aliphatic	Daphnid	Acute	32 [0.87, 1200]		-0.15	Estimated	in	[-1.61, 4.76] in		0.78	0.38	0.55	7(1)	no
2			COc	45.08	32200841	amine primary unreactive NH2-1 aliphatic alga	Alga	Acute	5.6 [0.0076, 4100]		-0.15	Estimated	in	[-1.61, 4.76] in		0.43	-0.61	1.02	7(0)	no
2			COc	45.08	12500251	CNO,X unreactive Fish Chronic, w/ N,O	Fish	Chronic	0.25 [0.012, 5.4]		-0.15	Estimated	in	[-1.61, 5.99] in		0.62	0.54	0.57	19(2)	no
2			COc	45.08	22200851	amine primary unreactive NH2-1 aliphatic	Daphnid	Chronic	0.69 [0.039, 12]		-0.15	Estimated	in	[-1.61, 1.63] in		0.23	-2.22	0.26	4(0)	no
2			COc	45.08	32200851	amine primary unreactive NH2-1 aliphatic alga	Alga	Chronic	0.86 [0.0039, 190]		-0.15	Estimated	in	[-1.61, 4.76] in		0.6	-0.08	0.84	7(0)	no
3			COd	46.07	10600041	CO,X primary alcohol	Fish	Acute	3700 [390, 34000]		-0.14	Estimated	in	[-1.75, 5.26] in		0.92	0.9	0.44	22(15)	yes
3			COd	46.07	12102041	CO,X alcohol unreactive w/o EO Fish	Fish	Acute	6200 [690, 54000]		-0.14	Estimated	in	[-0.63, 5.81] in		0.89	0.88	0.45	46(13)	yes
3			COd	46.07	12899941	narcotic group Fish Acute	Fish	Acute	3600 [480, 27000]		-0.14	Estimated	in	[-0.63, 5.88] in		0.87	0.87	0.43	151(34)	yes
3			COd	46.07	20600041	CO,X primary alcohol	Daphnid	Acute	1300 [150, 12000]		-0.14	Estimated	out(p)	[2.31, 5.26] in		0.95	0.76	0.17	6(17)	yes
3			COd	46.07	22102241	CO,X alcohol unreactive w/o EO Daphnid	Daphnid	Acute	750 [25, 22000]		-0.14	Estimated	out(p)	[0.78, 5.81] in		0.78	0.72	0.55	14(13)	yes
3			COd	46.07	22899941	narcotic group Daphnid Acute	Daphnid	Acute	240 [28, 2100]		-0.14	Estimated	out(p)	[1.08, 5.88] in		0.71	0.7	0.43	81(24)	yes
3			COd	46.07	30600041	CO,X primary alcohol	Alga	Acute	28000 [260, 2.9e+6]		-0.14	Estimated	out(p)	[2.31, 5.26] in		0.91	0.79	0.36	6(16)	yes
3			COd	46.07	32102041	CO,X alcohol unreactive w/o halogen, acid, EO	Alga	Acute	12000 [380, 360000]		-0.14	Estimated	out(p)	[1.08, 5.26] in		0.95	0.9	0.35	6(14)	yes
3			COd	46.07	32899941	narcotic group Alga Acute	Alga	Acute	4800 [420, 54000]		-0.14	Estimated	out(p)	[1.08, 5.26] in		0.75	0.73	0.44	50(47)	yes

図7-10 「複数の化学物質予測」結果タブ区切り (tsv) 形式ファイルの中身

- ※ 構造式の画像は含まれません。
- ※ criteria の列については、統計指標の基準 ($R^2 \geq 0.7$ and $Q^2 \geq 0.5$ and $n \geq 5$) を満たすものについては yes、他の場合は no と記載されています。
- ※ エラーが発生している物質は、QSAR ID 列に << Error >> とのみ記載し、エラー内容は出力されません。エラー内容についてはWeb側でご確認下さい。



計算に不具合があるSMILES等を見つけた場合は、KATE担当 (KATE@nies.go.jp) までご一報頂けますと幸いです。

8. 一括印刷フォーマット表示 (Print Format画面)

予測結果 (前述図 5-1) で、**Create Print Format** ボタンをクリックすると、Results画面および「Print Detail」列でチェックが入ったQSARクラスのVerify QSAR画面を一括して印刷するためのフォーマット画面が表示されます (図 8)。

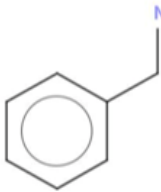
(1) → Ecotoxicity Prediction by KATE2025 version 1.0
March 7, 2025 at 17:07 (JST) [https://kate3.nies.go.jp]

(2) → **Results**

Summary of the Query Chemical

Query Chemical

SMILES	NCc1ccccc1	
CAS RN*		
Chemical Name		
log P	User Input Value	
	Estimated Value by KOWWIN™	1.07
	Measured Value in KOWWIN™ Database	1.09
Molecular Weight	107.15	



QSAR Prediction Result

Toxicity filter: all Fish Daphnid Alga
Acute Chronic

Statistical filter: Apply the following statistical criteria:
 R² ≥ 0.7 Q² ≥ 0.5 n ≥ 5

Print Detail	QSAR Class Name	Type of Predicted Toxicity		Predicted Toxicity [mg/L]	95% Prediction Interval	log P (Estimated)	Applicability Domain Judgement		Statistics of QSAR Class					
		Organism	Acute or Chronic				log P [Range]	Structure	R ²	Q ²	RMSE	n	criteria	
<input checked="" type="checkbox"/>	amine primary unreactive NH2=1 aliphatic	Fish	Acute	86	[6.2, 1200]	1.07	in	[-1.61, 5.25]	in	0.84	0.81	0.54	26(2)	✓
<input checked="" type="checkbox"/>	amine primary unreactive NH2=1 aliphatic	Daphnid	Acute	18	[0.54, 600]	1.07	in	[-1.61, 4.76]	in	0.78	0.38	0.55	7(1)	
<input checked="" type="checkbox"/>	amine primary unreactive NH2=1 aliphatic alga	Alga	Acute	4.1	[0.0064, 2600]	1.07	in	[-1.61, 4.76]	in	0.43	-0.61	1.02	7(0)	
<input type="checkbox"/>	CNO X unreactive Fish Chronic, w/ N.O	Fish	Chronic	0.22	[0.011, 4.1]	1.07	in	[-1.61, 5.99]	in	0.62	0.54	0.57	19(2)	
<input type="checkbox"/>	amine primary unreactive NH2=1 aliphatic	Daphnid	Chronic	1.2	[0.065, 24]	1.07	in	[-1.61, 1.63]	in	0.23	-2.22	0.26	4(0)	
<input type="checkbox"/>	amine primary unreactive NH2=1 aliphatic alga	Alga	Chronic	0.53	[0.0027, 110]	1.07	in	[-1.61, 4.76]	in	0.60	-0.08	0.84	7(0)	

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

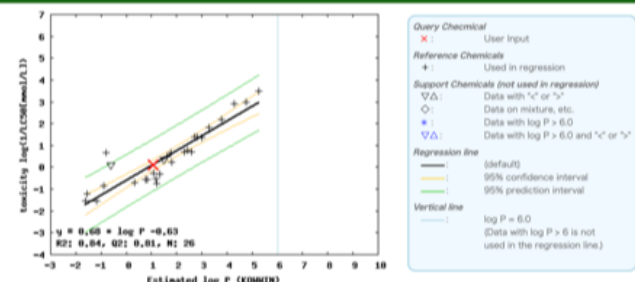


図 8 Print Format 画面

- (1) KATE2025のバージョンおよび予測結果が出力された日時 (日本標準時)
- (2) Results画面 (チェックは変更できない)
- (3) チェックが入ったQSARクラスに対するVerify QSAR画面の主な部分 (log(1/EC₅₀ 等[mmol/L]) vs log Pグラフ、予測対象物質の当該QSARクラスでの予測結果 (予測毒性値、95%予測区間、適用領域の判定等)、QSARクラスの回帰式と統計値、物質データ、構造判定用部分構造一覧等)

