

Toxicity Predictor  
(LIVER/MIE-QSAR & LUNG/MIE-QSAR)

ユーザガイド

2021 年 7 月 第 1.5 版

1. システム概要 .....	3
2. 推奨される動作環境 .....	3
3. 利用方法.....	4
3.1. ホーム画面 .....	4
3.1.1. 2次元構造エディタによる描画.....	4
3.1.2. SMILES による構造の入力.....	5
3.1.3. SDF ファイルのアップロード .....	7
3.1.4. SMILES リストのアップロード.....	7
3.2. 予測結果画面.....	8
3.2.1. 予測完了の待機 .....	8
3.2.2. 予測結果の確認 .....	8
3.2.3. 肝毒性/肺毒性予測結果の確認.....	10
3.2.4. MIE 予測結果の確認.....	12
3.2.5. 類似医薬品の確認.....	13
3.2.6. 予測結果のダウンロード .....	14
3.3. 予測モデルの性能確認画面.....	16
3.4. ユーザ登録/ログイン画面.....	17
謝辞 .....	18
Acknowledgements.....	18

## 1. システム概要

Toxicity Predictor は、入力された化合物に対して AOP における 59 種類の初期イベント MIE (Molecular Initiating Event) の活性を予測し、その予測値をもとに薬剤性肝毒性 (Drug Induced Liver Injury, DILI) と薬剤性肺毒性 (Drug Induced Pulmonary Toxicity, DIPT) に関する予測を行う定量的構造活性相関 (QSAR) システムです。予測モデルは PubChem の活性値データや、JAPIC AERS の提供する副作用症例報告データを訓練データとして機械学習を行うことで構築されており、予測値に加え、類縁医薬品に関する症例報告データも確認することが可能です。化合物は Web ブラウザ上で構造式を描画するほか、SMILES や SDF 形式で入力することができます。

## 2. 推奨される動作環境

本システムの使用に当たり、推奨される Web ブラウザは以下のとおりです。

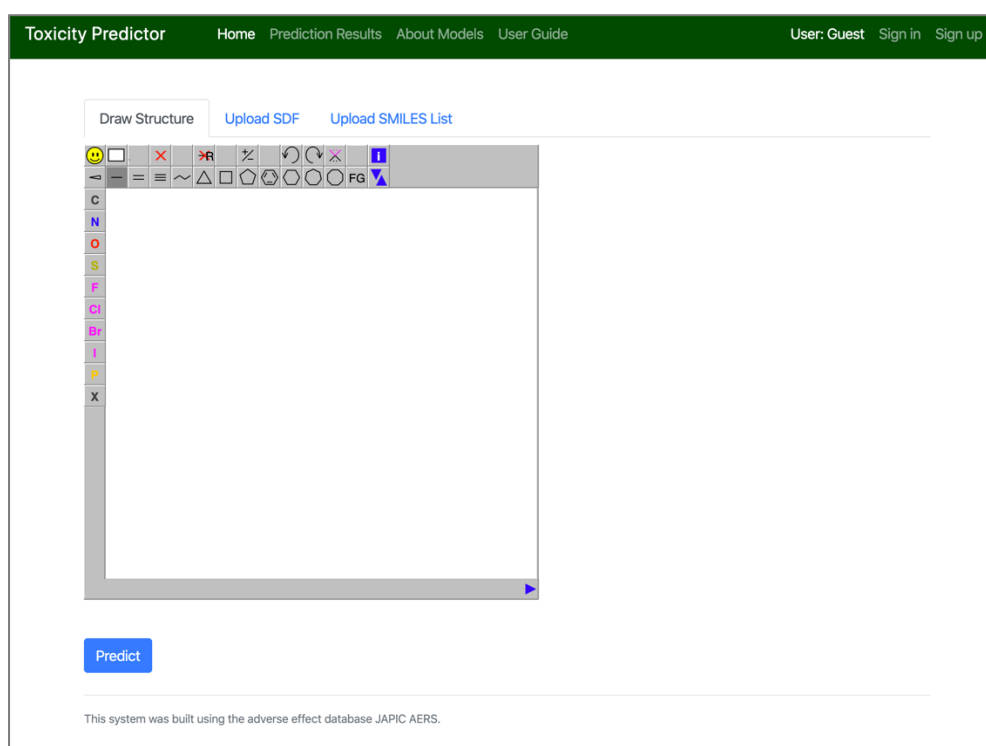
ブラウザ名	バージョン
Google Chrome	78.0 以降
Firefox	70.0 以降
Safari	12.0.3 以降
Microsoft Edge	44 以降
Internet Explorer	11 以降

### 3. 利用方法

本章では、Toxicity Predictor の利用方法について説明します。

#### 3.1. ホーム画面

Toxicity Predictor の URL (<http://mmi-03.my-pharm.ac.jp/tox1/>) にアクセスすると、以下のようなホーム画面が表示されます。ホーム画面からは予測対象の化合物を入力することが可能です。化合物の入力方法は、2次元構造エディタによる描画と、SMILES の入力、SDF のアップロード、SMILES リストのアップロードの4種類があります。



##### 3.1.1. 2次元構造エディタによる描画

2次元構造エディタを用いて化合物を入力する場合は、以下のように構造式をエディタ上で描画したあとに Predict ボタンをクリックします。Predict ボタンのクリック後、予測結果画面 (3.2 節) に遷移します。

Toxicity Predictor Home Prediction Results About Models User Guide User: Guest Sign in Sign up

Draw Structure Upload SDF Upload SMILES List

← ①構造式を描画

Predict ← ②Predict をクリック

This system was built using the adverse effect database JAPIC AERS.

### 3.1.2. SMILES による構造の入力

SMILES を使って化合物の毒性予測を行う場合は、①構造エディタ上で右クリックメニューを表示してから、②Paste MOL or SDF or SMILES をクリックし、③表示されたウィンドウ内に SMILES を入力した後、④Accept ボタンをクリックします。その後、構造エディタ上に入力した化合物の構造が表示されるため、確認して⑤Predict ボタンをクリックします。Predict ボタンのクリック後、予測結果画面（3.2 節）に遷移します。

なお、上記の方法で MOL 形式や SDF 形式を入力することも可能ですが、この場合 SDF 内の最初の 1 化合物のみに対して予測が行われます。SDF に複数の化合物が含まれている場合は後述する SDF のアップロードによって入力を行ってください（3.1.3 項）。また、複数の SMILES が含まれたファイルを用意してアップロードすることも可能です（3.1.4 項）。

Draw Structure

[Upload SDF](#)[Upload SMILES List](#)

The screenshot shows the drawing area of the Toxicity Predictor. A right-click context menu is open, listing several options. A red callout box with the number 1 points to the drawing area, and another red callout box with the number 2 points to the 'Paste MOL or SDF or SMILES' option in the menu.

① エディタ上で右クリック

- Copy as SMILES
- Copy as MOL
- Copy as MOL V3000
- Copy as InChI
- Copy as InChI key
- Search chemical structure (through InChIKey)
- Copy as Scalar Vector Graphics
- Paste MOL or SDF or SMILES

↑ ② Paste MOL or...を選択

Draw Structure

[Upload SDF](#)[Upload SMILES List](#)

The screenshot shows the drawing area with a 'Paste' dialog box open. The dialog box contains a text input field with the SMILES string 'c1ccccc1' and an 'Accept' button. A red callout box with the number 3 points to the input field, and another red callout box with the number 4 points to the 'Accept' button.

Paste

Paste the text to import into the text area below. Or drag and drop a file on it.

c1ccccc1

← ③ SMILES を入力

↑ ④ Accept をクリック

Accept ファイルを選択 選択されていません Cancel

Predict

↑ ⑤ Predict をクリック

### 3.1.3. SDF ファイルのアップロード

SDF ファイルをアップロードすることで、複数の化合物に関する毒性予測を同時に開始することができます。なお同時に予測を行える化合物の数は、ユーザごとに 100 件までに制限されています。SDF をアップロードするには、①Upload SDF をクリックしてタブを切り替えた後、②Select SDF をクリックしてアップロードする SDF ファイルを選択し、③Predict ボタンをクリックします。Predict ボタンのクリック後、予測結果画面 (3.2 節) に遷移します。

The screenshot shows the Toxicity Predictor web application interface. At the top, there is a navigation bar with 'Toxicity Predictor', 'Home', 'Prediction Results', 'About Models', and 'User Guide'. On the right, it says 'User: Guest' with 'Sign in' and 'Sign up' links. Below the navigation bar, there are three tabs: 'Draw Structure', 'Upload SDF', and 'Upload SMILES List'. The 'Upload SDF' tab is active, and a red callout box with a left-pointing arrow points to it, containing the text '①Upload SDF をクリック'. Below the tabs, there is a file input field containing '4\_compounds.sdf' and a blue 'Select SDF' button. A red callout box with an upward-pointing arrow points to the 'Select SDF' button, containing the text '②Select SDF をクリック ファイルを選択'. Below the file input field, there is a blue 'Predict' button. A red callout box with an upward-pointing arrow points to the 'Predict' button, containing the text '③Predict をクリック'.

### 3.1.4. SMILES リストのアップロード

SMILES を複数同時に入力したい場合は、SMILES のリストをアップロードします。アップロードするファイルには、各行に SMILES を 1 つのみ記述し、それ以外の情報は含めないようにしてください。同時に予測を行える化合物の数は、SDF の場合と同じくユーザごとに 100 件までに制限されています。SMILES リストをアップロードするには、①Upload SMILES List をクリックしてタブを切り替えた後、②Select SMILES List をクリックしてアップロードするファイルを選択し、③Predict ボタンをクリックします。Predict ボタンのクリック後、予測結果画面 (3.2 節) に遷移します。

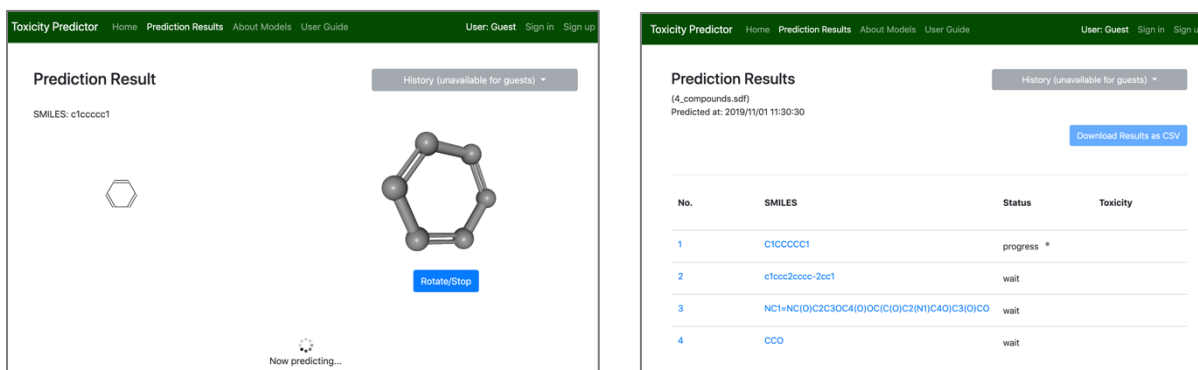
The screenshot shows the Toxicity Predictor web application interface. At the top, there is a navigation bar with 'Toxicity Predictor', 'Home', 'Prediction Results', 'About Models', and 'User Guide'. On the right, it says 'User: Guest' with 'Sign in' and 'Sign up' links. Below the navigation bar, there are three tabs: 'Draw Structure', 'Upload SDF', and 'Upload SMILES List'. The 'Upload SMILES List' tab is active, and a red callout box with a left-pointing arrow points to it, containing the text '①Upload SMILES List をクリック'. Below the tabs, there is a file input field containing '2\_compounds.smi' and a blue 'Select SMILES List' button. A red callout box with an upward-pointing arrow points to the 'Select SMILES List' button, containing the text '②Select SMILES List をクリック ファイルを選択'. Below the file input field, there is a blue 'Predict' button. A red callout box with an upward-pointing arrow points to the 'Predict' button, containing the text '③Predict をクリック'. Below the file input field, there is a message: 'Uploaded file should contain one SMILES for each line'.

## 3.2. 予測結果画面

予測結果画面では、毒性予測の結果を確認することが可能です。本画面には、予測を開始した際に自動的に遷移するほか、上部メニューの「Prediction Results」をクリックすることで遷移することができます。

### 3.2.1. 予測完了の待機

毒性予測を開始すると、以下のような待機画面に移動します。入力した化合物が単一の化合物か複数の化合物かによって表示される画面は異なります（単一の化合物の場合は下図左のような画面、複数の化合物の場合は下図右のような画面が表示されます）。それぞれの化合物の予測には、おおよそ1分程度かかりますが、この時間はサーバの混雑状況や化合物サイズなどによって変動します。



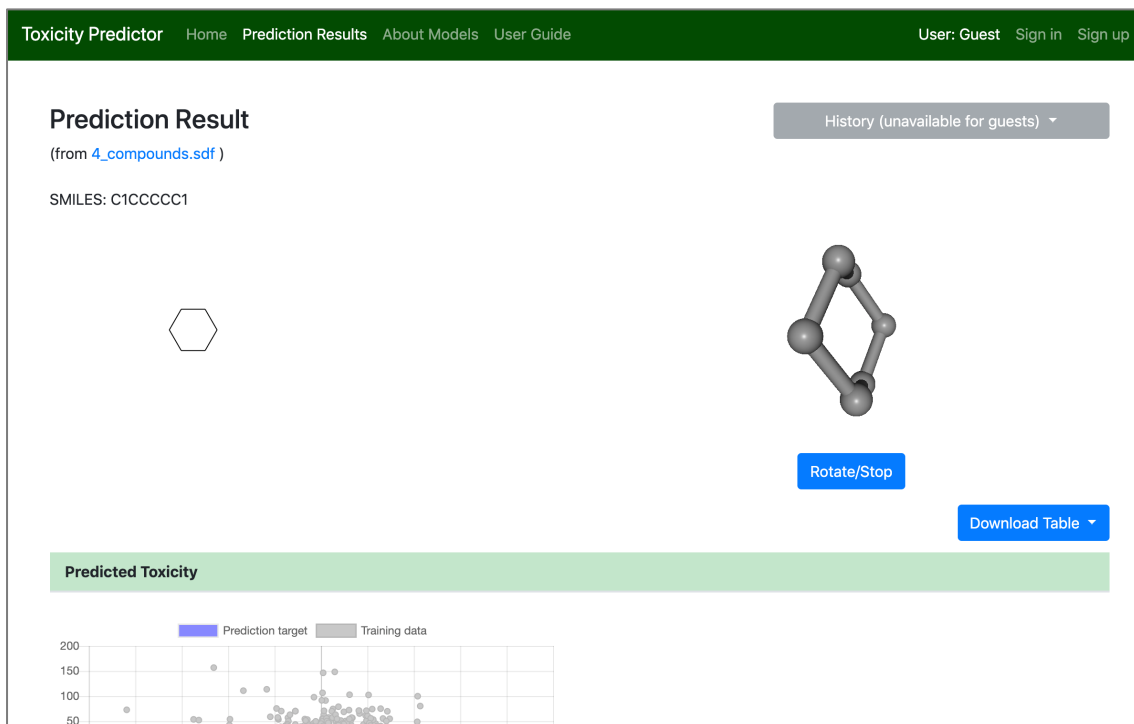
The left screenshot shows the 'Prediction Result' page for a single compound. The SMILES string is c1ccccc1. It displays a 2D skeletal structure of benzene and a 3D ball-and-stick model. A 'Rotate/Stop' button is located below the 3D model. The status at the bottom indicates 'Now predicting...'. The right screenshot shows the 'Prediction Results' page for multiple compounds. It includes a 'Download Results as CSV' button and a table with the following data:

No.	SMILES	Status	Toxicity
1	<chem>C1CCCCC1</chem>	progress *	
2	<chem>c1ccc2ccccc2cc1</chem>	wait	
3	<chem>NC1=NC(O)C2C3OC4(O)OC(C(O)C2)N1C4(O)C3(O)CO</chem>	wait	
4	<chem>CCO</chem>	wait	

### 3.2.2. 予測結果の確認

予測が完了すると、予測結果が表示されます。単一の化合物の予測を行った場合は、以下のように完了した時点で予測結果詳細が表示されている状態になります。本画面の詳細に関しては、3.2.3 項以降を参照してください。





複数の化合物の予測を行った場合は、以下のように化合物ごとに予測の進捗状況が表示されます。予測が完了すると、**Status** 列が **success** になります。予測が完了した化合物に関しては、**No.**列もしくは **SMILES** 列をクリックすることで化合物ごとの予測結果詳細を閲覧できます。また、すべての化合物の予測が完了すると、予測結果の一覧を **CSV** でダウンロードすることができるようになります (**CSV** の各列の意味は **3.2.6** 節を参照してください)。また、ユーザ登録およびログイン (**3.4** 節) をしている場合は右上のボタンからこれまでに予測した化合物の履歴を閲覧できます。

Toxicity Predictor Home Prediction Results About Models User Guide

↓ 履歴の閲覧 (要ログイン)

### Prediction Results



(4\_compounds.sdf)  
Predicted at: 2019/11/01 11:30:30

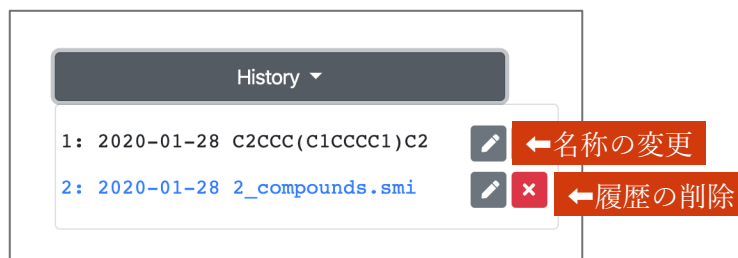
Download Results as CSV

↑ CSV によるダウンロード

No.	SMILES	Status	
1	<a href="#">C1CCCCC1</a>	success	equivocal
2	<a href="#">c1ccc2cccc-2cc1</a>	success	positive
3	<a href="#">NC1=NC(O)C2C3OC4(O)OC(C(O)C2(N1)C4O)C3(O)CO</a>	success	positive
4	<a href="#">CCO</a>	success	positive

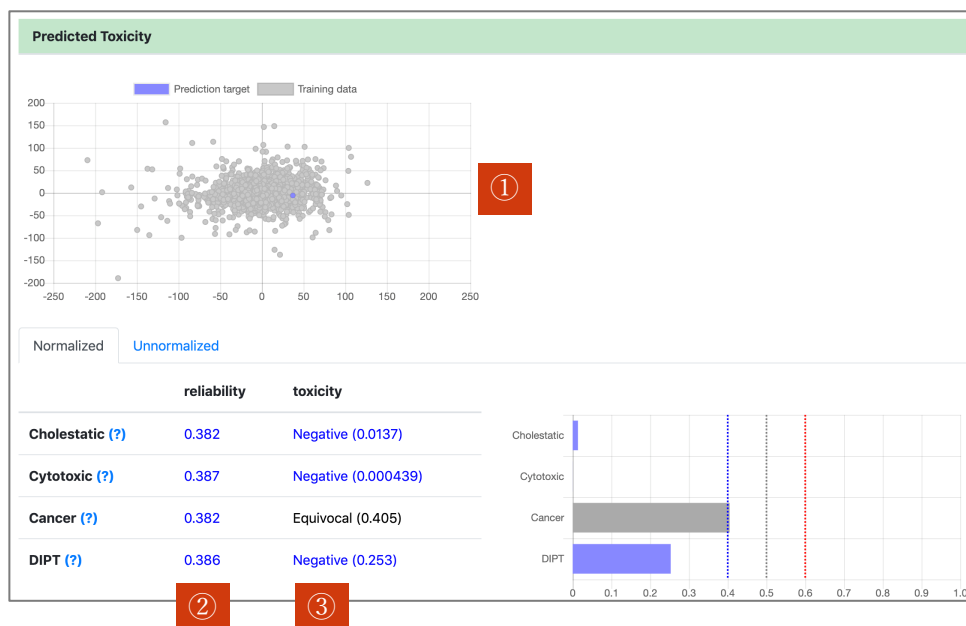
↑ 個別の予測結果へのリンク

履歴ボタンをクリックすると、以下のようにそれまでに予測したファイル名や SMILES の履歴が表示されます。このウィンドウ内で、 ボタンをクリックするとその予測履歴の名称を自由に変更することができます。また、 ボタンをクリックすることでその履歴を削除することができます。



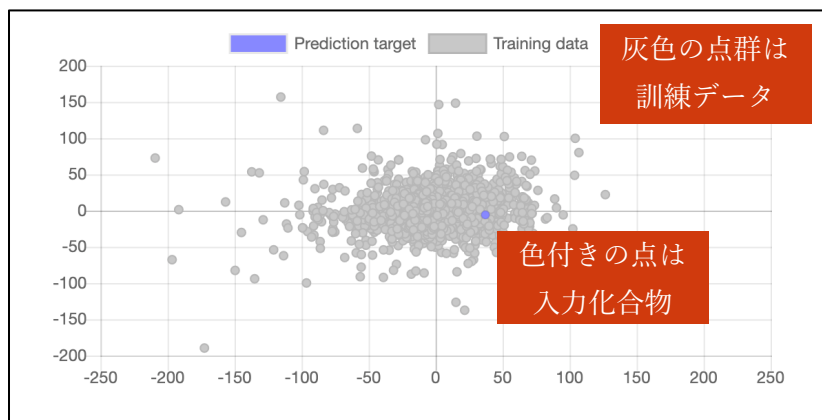
### 3.2.3. 肝毒性/肺毒性予測結果の確認

予測結果詳細の画面では、まず以下のように肝毒性と肺毒性の予測結果が表示されます。表示される内容は、それぞれ①訓練データの散布図、②毒性予測の信頼度、③毒性の予測値です。



①訓練データの散布図は、入力化合物が訓練データに対する大まかな位置関係を図示したものです。本散布図中で、訓練データは灰色の点群で表示され、予測対象の化合物は色付きの点で表示されています。各点の位置は学習に使用された化合物データセットから得られた化学構造記述子の組をそれぞれ高次元ユークリッド空間中の点と見なし、それらの点の集合を二次元ユークリッド空間に圧縮することで計算されています。入力化合物の色は②毒性予測の信頼度の値によって、**緑** (既知化合物の場合)、**青** (信頼度：高)、**赤** (信頼度：低) に色分けされま

す。この信頼度は、圧縮前の高次元ユークリッド空間における、入力化合物とその近傍の化合物との距離の相乗平均を元に算出されています。



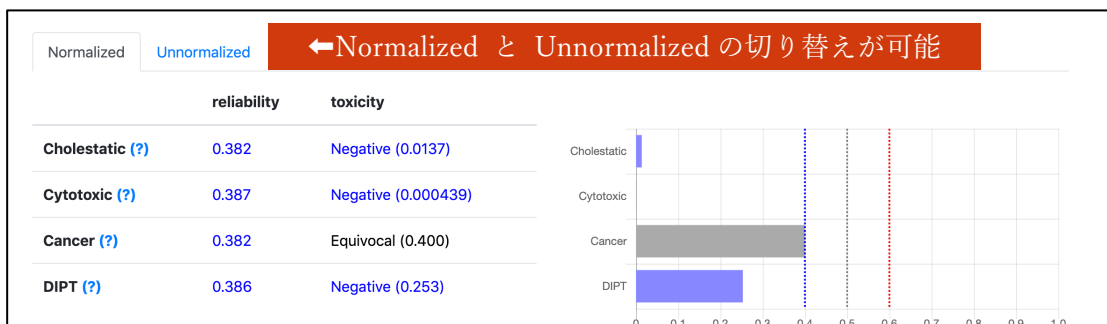
③毒性の予測値は、数値と棒グラフで表示されます。予測対象となる毒性は以下の通りです。また、毒性名の右に位置する「(?)」のリンクをクリックすることで、対応する予測モデルの性能 (3.3 節) を確認することが可能です。

表示名	意味
Cholestatic	胆汁うっ滞性肝障害
Cytotoxic	細胞傷害性肝障害
Cancer	肝臓がん
DIPT	薬物誘発性肺毒性

毒性表示のオプションとしては Normalized タブと Unnormalized タブが用意されており、デフォルトでは Normalized タブが選択されています。Normalized タブ内で表示されている各予測値 $x_n$ は、予測モデルの直接の出力値  $x_u$  を元に、以下の式による正規化をおこなったものになっています ( $c$ は各予測モデルにおける ROC 曲線上の Youden Index に基づいたカットオフを意味します)。

$$x_n := x_u^{-\log_c 2}$$

Unnormalized タブをクリックすることで正規化前の $x_u$ の値を確認することができます。



本表内では、 $x_n$ の値によって色分けを行っており、0.4未満がNegative (青)、0.4以上0.6未満がEquivocal (黒)、0.6以上がPositive (赤)と表示されます。なお、入力された化合物がJAPIC AERSにおける毒性既知の化合物であった場合には、毒性値は予測値ではなく実験値が表示されます。化合物が既知化合物かどうかを判定する際には、入力したSMILESが「@」、「/」、「¥」のいずれかを含む場合のみ立体異性体を区別し、それ以外の場合は2次元構造を使用して照合します。

### 3.2.4. MIE 予測結果の確認

肝毒性/肺毒性予測結果部分の下には、以下のように59種類のMIEの予測値が表示されます。基本的な図表の意味は肝毒性/肺毒性予測の部分と同様ですが、MIEに関してはPubChem上の活性値スコア $s$ に対し、 $s \geq 1$ となるものを陽性とみなして訓練した予測モデルと、 $s \geq 40$ となるものを陽性とみなして訓練した予測モデルの2種類を用意しているため、1つのMIEに関してそれぞれ2つずつの予測値を表示しています。また、一部の十分な予測性能が得られなかったモデルに関しては、「Criteria {1, 40} is unavailable」のように表示されます。



### 3.2.5. 類似医薬品の確認

MIE 予測結果部分の下には、入力された化合物に対し、JAPIC AERS の症例報告から取得した類似医薬品が表示されます。

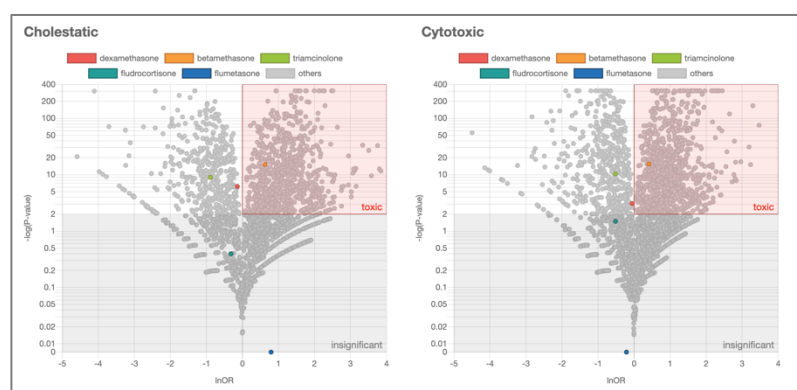
Similar Medicines												
name	similarity	n	Cholestatic			Cytotoxic			Cancer			DIPT
			toxicity	ROR	p-value	toxicity	ROR	p-value	toxicity	ROR	p-value	
calcium	0.333	566028	Negative	0.371	1.41×10 <sup>-144</sup>	Negative	0.457	1.63×10 <sup>-276</sup>	Negative	0.540	4.88×10 <sup>-12</sup>	Nega
hydrochloric acid	0.333	1698	Negative	0.663	0.600	Negative	1.02	1.00	Negative	2.64	0.434	Nega
oxaliplatin	0.182	138486	Positive	1.78	1.53×10 <sup>-34</sup>	Positive	2.56	1.11×10 <sup>-302</sup>	Positive	1.54	0.000682	Positi
sodium perchlorate	0.167	1028	Positive	3.30	0.00894	Positive	2.35	0.00677	Positive	7.28	0.0472	Nega
cyclamic acid	0.158	143	Negative	1.57	1.00	Negative	2.90	0.217	Negative	10.4	1.00	Nega

各類似医薬品に関し、以下の情報が表示されます。

表示名	意味
similarity	入力化合物に対する類似の度合い (0 ~ 1)
n	症例報告の報告件数

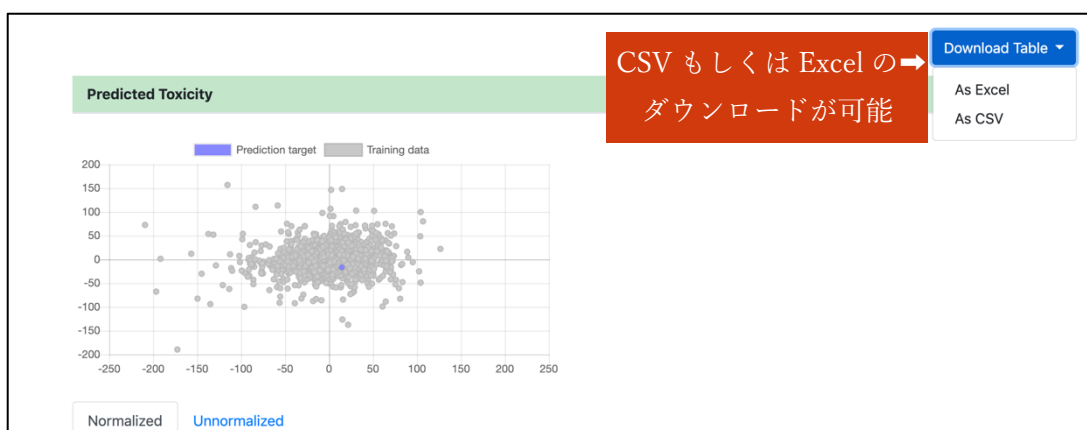
toxicity	毒性の有無 (Negative, Positive) ただし、症例報告の P 値が水準に満たない場合は Insignificant
ROR	症例報告オッズ比
p-value	症例報告の有意性を表す P 値
canonical smiles	類似医薬品の化学構造を表す SMILES
isomeric smiles	立体異性体の区別を考慮した SMILES

また、それぞれの毒性に関し、類似医薬品をハイライトしたボルケーノプロットが表示されます。ボルケーノプロット内では、オッズ比の値と P 値を元に、毒性を持つとみなされる領域が赤い領域で、P 値が基準に満たない領域が灰色の領域で表示されます。



### 3.2.6. 予測結果のダウンロード

画面右上部の「Download Table」より、予測結果を Excel ファイルもしくは CSV ファイルでダウンロードすることが可能です。



ダウンロードできる CSV の各列の意味は、以下の通りになっています。

列名	意味
SMILES	入力化合物を表す SMILES
<毒性名>_normalized	各種毒性の正規化後の毒性予測値
<毒性名>_unnormalized	各種毒性の正規化前の毒性予測値
<毒性名>_reliability	各種毒性の予測信頼度
<MIE 名>_normalized_{1, 40}	各種 MIE の正規化後の毒性予測値 1 および 40 は Positive とみなすための閾値を表す
<MIE 名>_unnormalized_{1, 40}	各種 MIE の正規化前の毒性予測値 1 および 40 の意味は上と同じ
<MIE 名>_reliability	各種 MIE の予測信頼度

### 3.3. 予測モデルの性能確認画面

上部メニューの「About Models」をクリックすることで、本システムが使用する各予測モデルの性能が確認できます。

Performance of Prediction Models										
Hepatotoxicity Prediction Models										
Type	AUC	Sensitivity	Specificity	Cutoff						
Cholestatic Liver Injury	0.807	0.735	0.768	0.831						
Cytotoxic Liver Injury	0.836	0.765	0.764	0.752						
Liver Cancer	0.815	0.736	0.754	0.161						
Pulmonary toxicity Prediction Models										
Type	AUC	Sensitivity	Specificity	Cutoff						
Drug-Induced Pulmonary Toxicity	0.850	0.748	0.842	0.134						
MIE Prediction Models										
Index	Description	AID	Criteria: 1				Criteria: 40			
			AUC	Sensitivity	Specificity	Cutoff	AUC	Sensitivity	Specificity	Cutoff
1	ATAD5_ind (ATAD5 genotoxic inducer)	720516 <a href="#">(PubChem)</a>	0.845	0.744	0.847	0.0370	0.840	0.750	0.843	0.0280
2	p53_ago (p53 agonist)	720552 <a href="#">(PubChem)</a>	0.845	0.804	0.793	0.142	0.899	0.824	0.830	0.0269
3	MMP_disr (mitochondrial membrane potential disruptor)	720637 <a href="#">(PubChem)</a>	0.795	0.698	0.788	0.368	0.919	0.845	0.846	0.0635

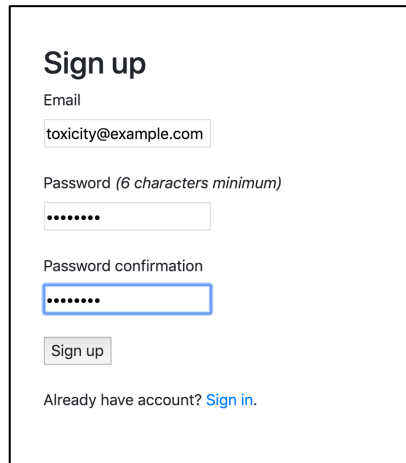
本画面で表示される内容は、以下のとおりです。

表示名	意味
Type	毒性名（肝毒性/肺毒性のみ）
AUC	評価データを用いて計算した ROC 曲線上の Area Under Curve の値
Sensitivity	予測モデルの感度 (TP / (TP + FN))
Specificity	予測モデルの特異度 (TN / (TN + FP))
Cutoff	ROC 曲線上の Youden Index に基づく Cutoff の値
Index	システム内での MIE の通し番号
Description	MIE の省略名および詳細名
AID	PubChem におけるアッセイ番号およびリンク



### 3.4. ユーザ登録/ログイン画面


本システムはユーザ登録を行わなくても使用することができますが、ユーザ登録およびログインを行うことで過去の毒性予測の履歴を確認することが可能になります (3.2 節)。ユーザ登録を行う場合は、まず画面上部メニューの「Sign up」をクリックしてから、登録するメールアドレスとパスワードを入力し、「Sign up」ボタンをクリックします。ユーザ登録が完了すると同時に、ログインが行われます。



The image shows a 'Sign up' form with the following fields and elements:

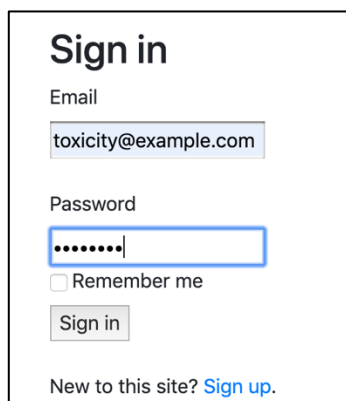
- Sign up** (Section Header)
- Email** (Label): toxicity@example.com
- Password (6 characters minimum)** (Label): [Redacted]
- Password confirmation** (Label): [Redacted]
- Sign up** (Button)
- Already have account? [Sign in.](#)** (Text)

ログイン中は、画面右上にユーザ名が表示されます。ログアウトしたい場合は、上部メニューの「Sign out」をクリックすることでログアウトできます。



The image shows a dark green header bar with the text: **User: toxicity Sign out**

再びログインを行いたい場合は、上部メニューの「Sign in」をクリック後、登録したメールアドレスとパスワードを入力してログインを行ってください。



The image shows a 'Sign in' form with the following fields and elements:

- Sign in** (Section Header)
- Email** (Label): toxicity@example.com
- Password** (Label): [Redacted]
- Remember me
- Sign in** (Button)
- New to this site? [Sign up.](#)** (Text)

## 謝辞

本システムの作成には、副作用症例報告データベースである JAPIC AERS のデータを使用しました。

## Acknowledgements

This system was built using the adverse effect database JAPIC AERS.