

## Table B – Use ToxCreate to build a (Q)SAR model

**Description of Activity:** Based on provided datasets build and examine a QSAR model.  
You can also try your own dataset.

**Requirements:** Browser with Java script activated (we recommend Firefox 5.0), Java installed, Internet connection.

**Duration for a Participant:** 45 minutes

**OpenTox Participants:** Christoph Helma, David Vorgrimmler

### Instructions:

#### 1 Create

- Open your browser and go to <http://www.toxcreate.org/> , <http://ot-test.in-silico.ch/toxcreate/> , <http://ot-dev.in-silico.ch/toxcreate/> or <http://mr.in-silico.ch/toxcreate/>. Follow the instructions on the webpage. Click on bold and purple topics to get further information.
- Upload training data in Excel, CSV or SDF format.
- After clicking on “Create model” it will take a while until the dataset has been uploaded
- If you want to use your own data follow the “instructions for creating training datasets” at <http://www.toxcreate.org/help/>)

Creates computational models to predict toxicity

**ToxCreate**

Create
Inspect
Predict
Help

You will need to upload training data that includes chemical structures and their measured toxicity values, in Excel , CSV or SDF file formats to create a prediction model. Please read the [instructions for creating training datasets](#) before submitting.

Upload training data in Excel , CSV or SDF format:  Browse...

Create model

This service creates and validates new *classification* and *regression* structure-activity models from your experimental data. The models can be used to predict toxicity of new chemicals (e.g. for REACH<sup>®</sup> purposes) and to reduce the need for animal testing. The following methods are currently available:

- *lazar classification* models and
- *lazar regression* models (experimental)

Further modelling algorithms may be added in future versions.

Disclaimer: ToxCreate uses state-of-the-art published and tested algorithms and methodologies with full validation information. However, just as with experimental measurements, computational predictions are subject to varying degrees of accuracy and uncertainty, so please read the full report carefully, particularly the validation information. No liability is accepted for any inaccuracy in predictions.

Version: v2.1.0 , Date: Thu Aug 4 18:38:58 2011 +0200 Date: Thu Aug 4 18:38:58 2011 +0200

© *in silico toxicology* 2009-2011, powered by **OpenTox** (a project funded by the **7th Framework Programme** of the European Commission)

## 2 Inspect

- On this page you get a brief summary of all models with validation results
- Find your model by name and click on bold and purple links to get/download detailed information (e.g. the feature dataset as XML or a detailed validation report)

Creates computational models to predict toxicity

**ToxCreate**

[Create](#) [Inspect](#) [Predict](#) [Help](#)

Get an overview about ToxCreate models. Parts of this page are refreshed every 5 seconds to update the model status.

### Fish\_Toxicity\_LC50\_mmol\_regression ( edit )

<b>Status:</b>	Completed( <a href="#">delete</a> )
<b>Started:</b>	08/05/2011 - 09:55:15AM GMT
<b>Training compounds:</b>	569
<b>Task:</b>	<a href="http://toxcreate3.in-silico.ch/task/447815">http://toxcreate3.in-silico.ch/task/447815</a> <a href="#">↗</a>
<b>Algorithm:</b>	lazar
<b>Type:</b>	regression
<b>Descriptors:</b>	Fminer backbone refinement classes <a href="#">↗</a>
<b>Training dataset:</b>	Excel sheet , SDF , YAML ( <i>experts</i> )
<b>Feature dataset:</b>	Excel sheet , SDF , YAML ( <i>experts</i> )
<b>Model:</b>	QMRF Editor, YAML ( <i>experts, models cannot be represented in Excel</i> )
<b>Validation:</b>	
<b>Detailed report:</b>	<a href="#">show</a>
<b>Number of predictions:</b>	452
<b>R-squared <a href="#">↗</a></b>	-0.00916
<b>Root Mean Square Error <a href="#">↗</a></b>	28
<b>Mean Absolute Error <a href="#">↗</a></b>	3.35

### ISSCAN\_v3a\_canc ( edit )

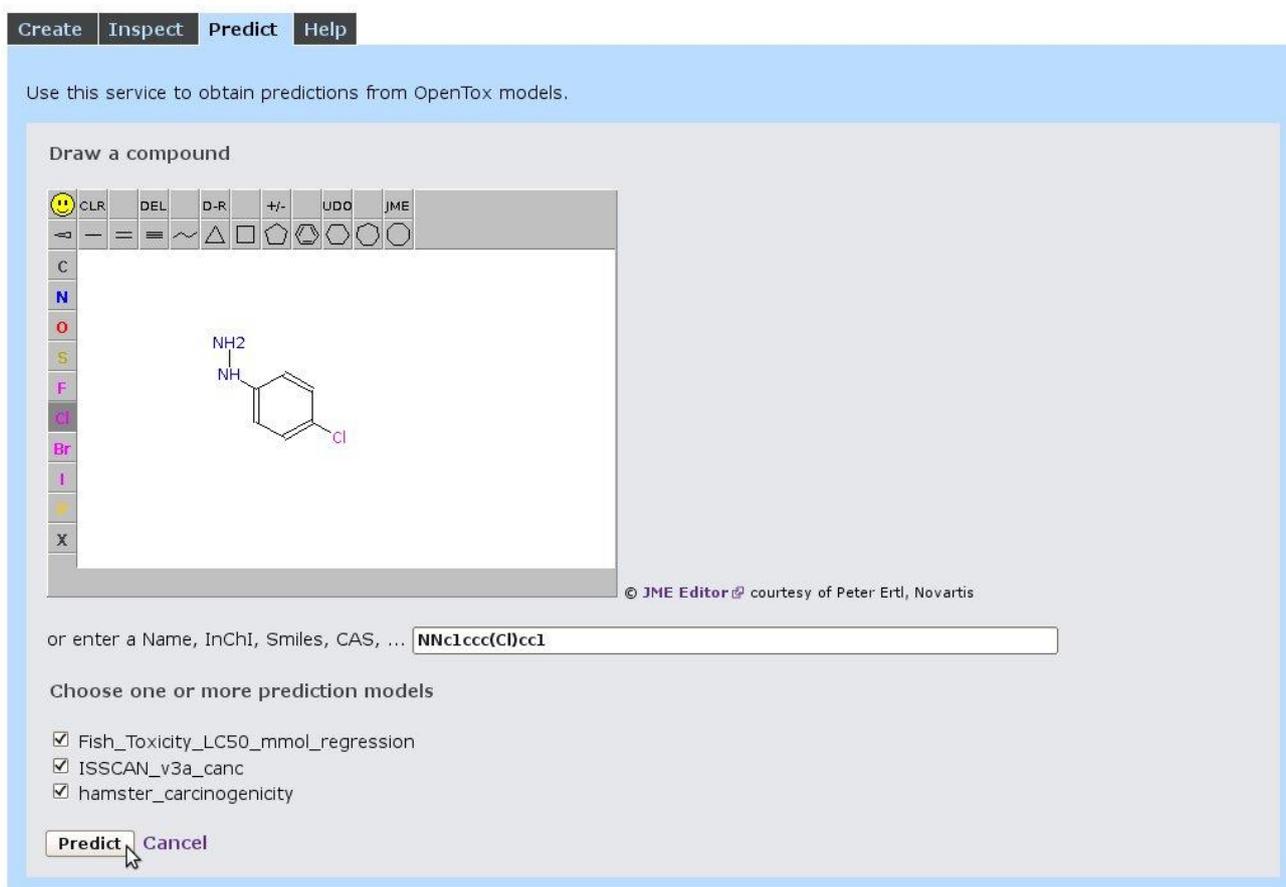
<b>Status:</b>	Completed( <a href="#">delete</a> )
<b>Started:</b>	08/05/2011 - 07:51:28AM GMT
<b>Training compounds:</b>	1056
<b>Task:</b>	<a href="http://toxcreate3.in-silico.ch/task/446614">http://toxcreate3.in-silico.ch/task/446614</a> <a href="#">↗</a>
<b>Warnings:</b>	<a href="#">show</a>
<b>Algorithm:</b>	lazar
<b>Type:</b>	classification

### 3 Predict

- Draw a compound or enter a Name, InChI, Smiles, CAS, ...
- Selected one or more prediction models and click on “Predict”

Creates computational models to predict toxicity

**ToxCreate**



Use this service to obtain predictions from OpenTox models.

Draw a compound

© JME Editor  courtesy of Peter Ertl, Novartis

or enter a Name, InChI, Smiles, CAS, ...

Choose one or more prediction models

- Fish\_Toxicity\_LC50\_mmol\_regression
- ISSCAN\_v3a\_canc
- hamster\_cardinogenicity

**Predict** Cancel

Disclaimer: ToxCreate uses state-of-the-art published and tested algorithms and methodologies with full validation information. However, just as with experimental measurements, computational predictions are subject to varying degrees of accuracy and uncertainty, so please read the full report carefully, particularly the validation information. No liability is accepted for any inaccuracy in predictions.

Version: v2.1.0 , Date: Thu Aug 4 18:38:58 2011 +0200 Date: Thu Aug 4 18:38:58 2011 +0200

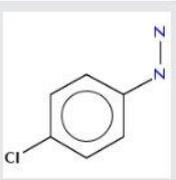
- Check the result and click on “Details” for prediction details (similar compounds, relevant substructures, ...)
- Click on bold and purple topics to get further information.

Creates computational models to predict toxicity

**ToxCreate**

[Create](#) [Inspect](#) [Predict](#) [Help](#)

New prediction

NNc1ccc(Cl)cc1			
	hamster_carcinogenicity:	ISSCAN_v3a_canc:	Fish_Toxicity_LC50_mmol_regression:
	0	1	0.0497
	( Confidence : 0.318 )	( Confidence : 0.178 )	( Confidence : 0.292 )
	<a href="#">Details</a>	<a href="#">Details</a>	<a href="#">Details</a>

Disclaimer: ToxCreate uses state-of-the-art published and tested algorithms and methodologies with full validation information. However, just as with experimental measurements, computational predictions are subject to varying degrees of accuracy and uncertainty, so please read the full report carefully, particularly the validation information. No liability is accepted for any inaccuracy in predictions.

Version: v2.1.0 , Date: Thu Aug 4 18:38:58 2011 +0200 Date: Thu Aug 4 18:38:58 2011 +0200

© in silico toxicology 2009-2011, powered by OpenTox (a project funded by the 7th Framework Programme of the European Commission)

- Detailed results of prediction are showing neighbors with relevant substructures, measured activity and similarity.
- Click on bold and purple topics to get further information.

Creates computational models to predict toxicity

**ToxCreate**

Create Inspect **Predict** Help

New prediction

Lazar prediction ( hide )

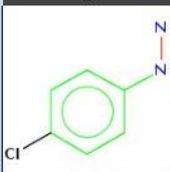
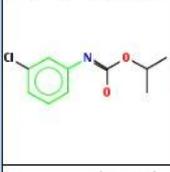
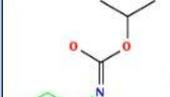
Lazar searches the training dataset for **similar** compounds (*neighbors*) and calculates the prediction from their measured activities. Lazar calculates predictions using

- a majority vote (weighted by compound similarity) for *classification* ( [original publication](#) )
- a local QSAR model based on neighbors for *regression* ( [original publication](#) )

Significant fragments are highlighted in the structure display as follows:

- features that occur predominately in compounds with activity: "1"
- features that occur predominately in compounds with activity: "0"
- regions, where fragments from different classes overlap
- inert parts

Please keep in mind that predictions are based on the measured activities of neighbors. Significant fragments are solely used to determine **activity specific similarities** of neighbors.

1	Prediction	Confidence	Supporting information
		0.318	Names and synonyms Significant fragments
Neighbors (1-5/15) next	Measured activity	Similarity	Supporting information
		0.875	Names and synonyms Significant fragments
		0.875	Names and synonyms Significant fragments