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.

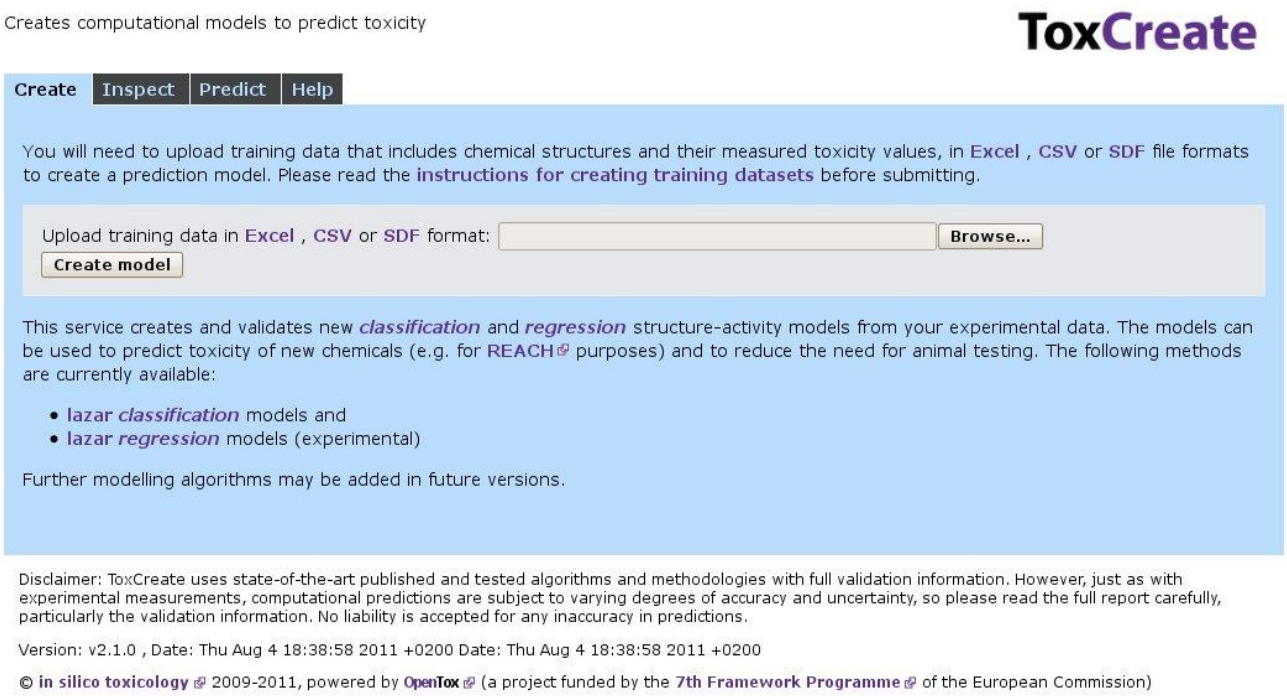
Duration for a Participant: 45 minutes

OpenTox Participants: Christoph Helma, David Vorgrimmler

Instructions:

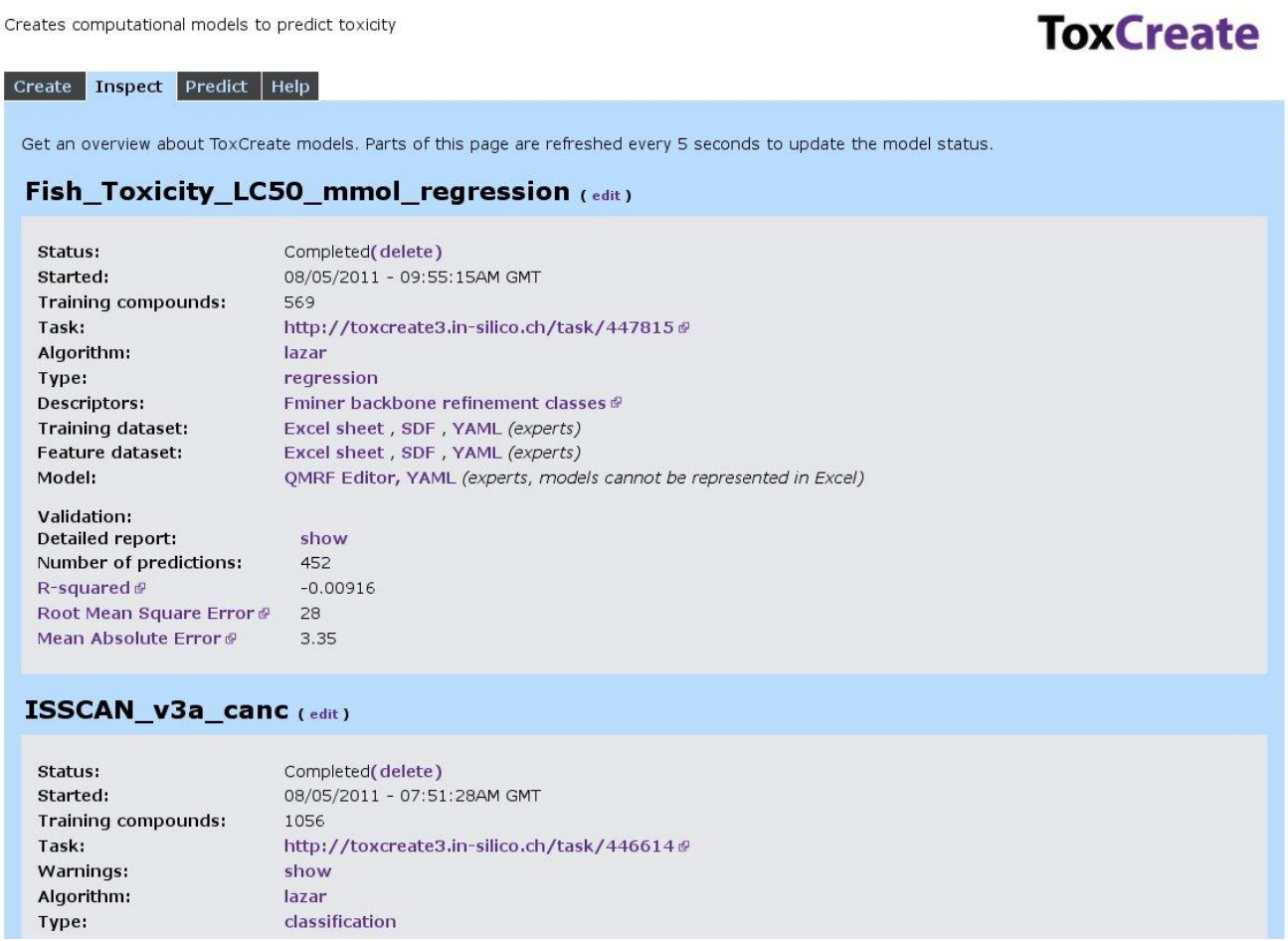
1 Create

- 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/)



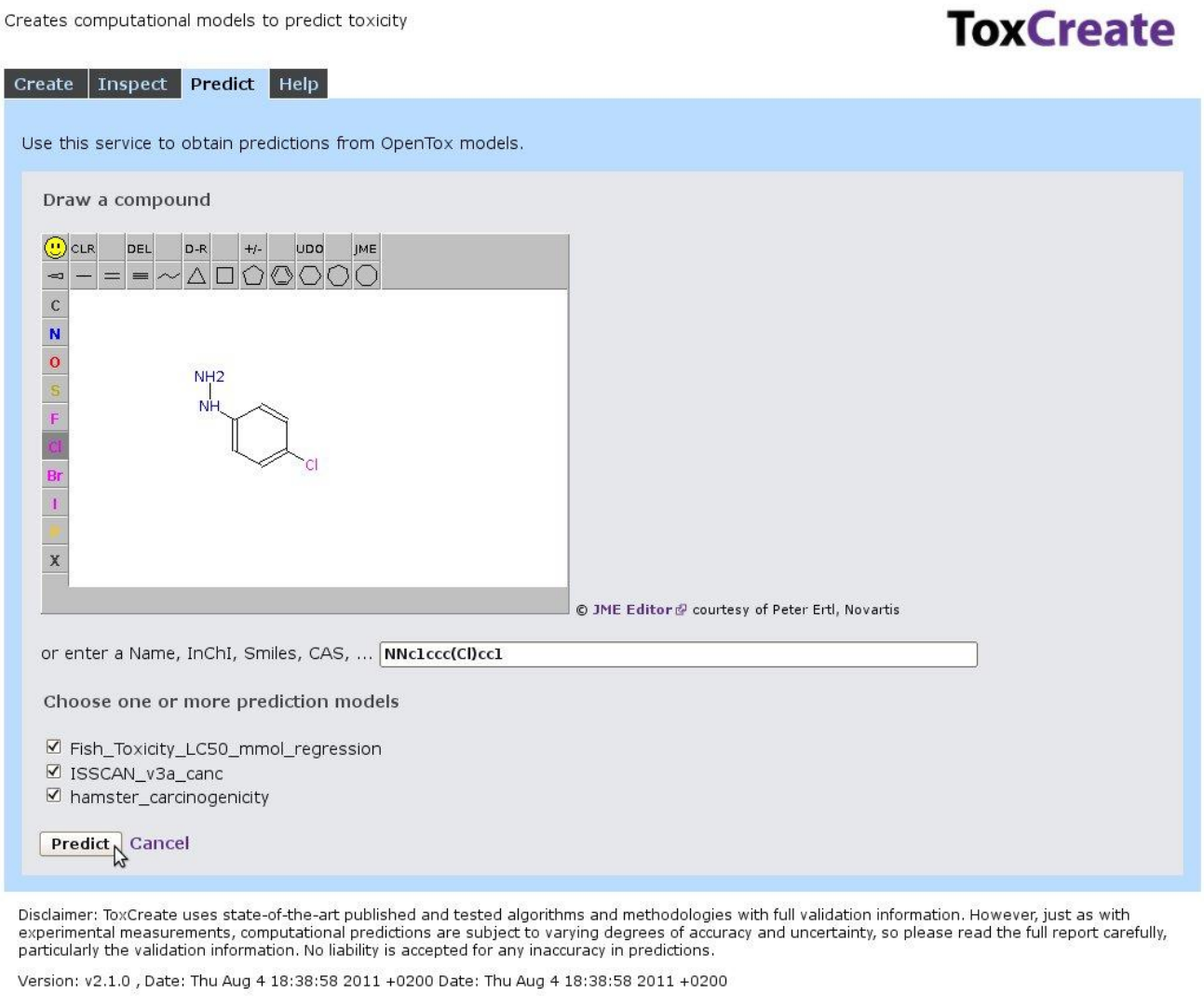
2 Inspect

- 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)

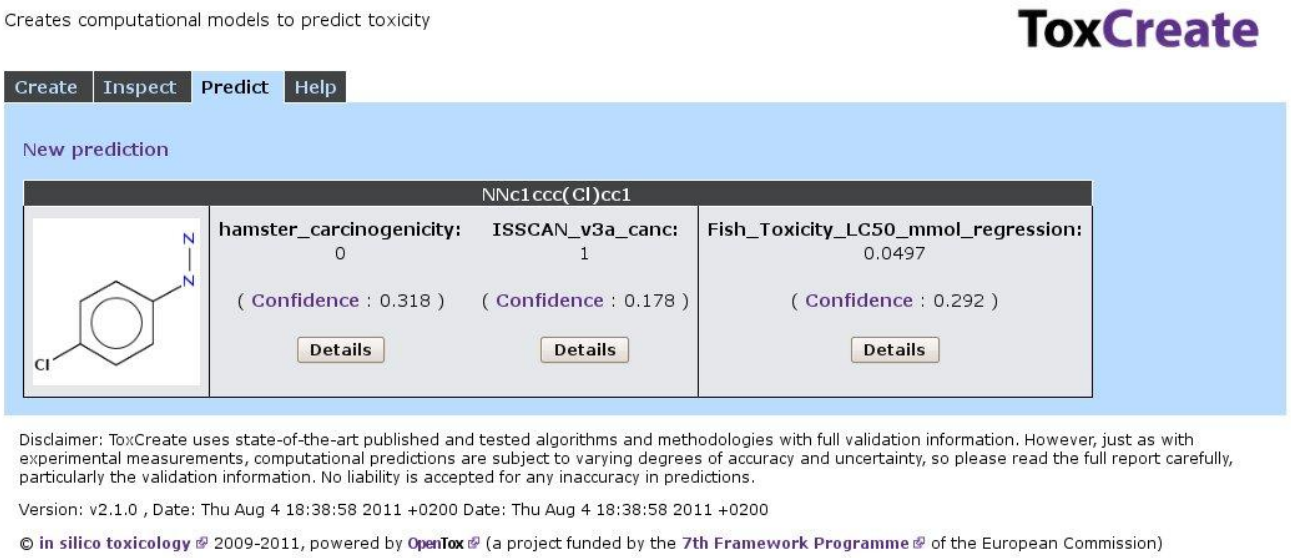


3 Predict

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



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



- 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)

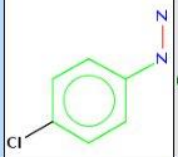
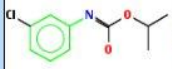
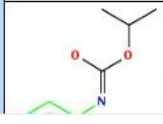
lazar searches the training dataset for **similar** compounds *(neighbors)* and calculates the prediction from their measured activities. lazar calculates predictions using

- 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	0.318	Names and synonyms Significant fragments
Neighbors (1-5/15) next	**Measured activity**	**Similarity**	**Supporting information**
	0	0.875	Names and synonyms Significant fragments
	0	0.875	Names and synonyms Significant fragments