DevTox-iS™

iS Tox Platform™
Machine-learning models for a new toxicology

Report

Client:
Altox

Username:
Tiago

Study Number:
DevTox-iS_Compound9_

Date:
2019/06/26 - 18:00:29

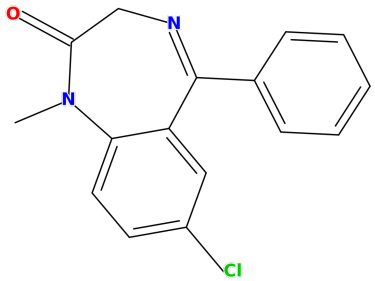
Program Version: 2.0

Molecular Query

Name:
Compound 9

CAS:
NA

SMILES:
CN1c2ccc(Cl)cc2C(=NCC1=O)c1ccccc1



Model Summary

DevTox-iS™ is a computational tool for prediction of the developmental toxicology by statistical and machine learning models. Potential *in vivo* toxicants for the developmental or reproductive endpoints and the active/inactive binders, agonists, and antagonists to the *in vitro* estrogen receptor (ER). The models were validated following the OECD (Organisation for Economic Co-operation and Development) Principles for the Validation for Regulatory Purposes of (Q)SAR Models. These OECD principles are discussed in each section of this report.

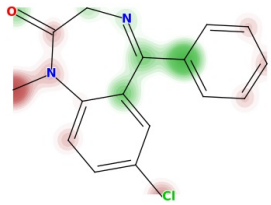
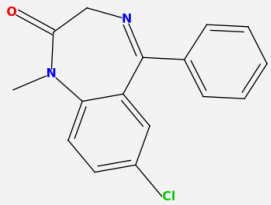
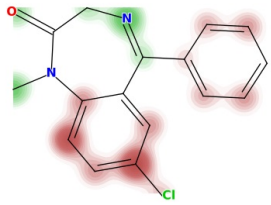
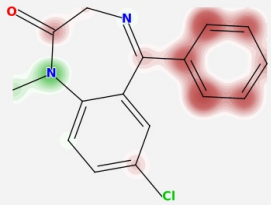
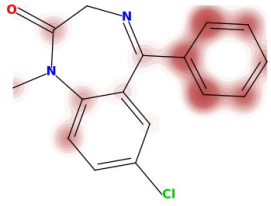
This tool is powered by artificial intelligence models for the Potential *in vivo* toxicants for the developmental or reproductive using 668 and 32,457 rigorously curated data of the categorized compounds into active/inactive binders, agonists, and antagonists to the *in vitro* estrogen receptor (ER) (defined endpoint - OECD principle 1).

The final result is provided in a summarising table with all the individual predictions, the applicability domain (AD), and relevant data for each endpoint/method.

Artificial Intelligence Models

The individual predictions below were obtained for the *in vivo* potential toxicity for the developmental and reproductive and, binders, agonists, and antagonists to the *in vitro* estrogen receptor (ER) by Artificial Neural Networks (ANN). To ensure transparency in the description of the model (an unambiguous algorithm - OECD Principle 2), more detailed information about each model is presented below.

The QSTR (quantitative structure-toxicity relationship) Probability Mapping indicates the fragments that decrease Toxicity / Activity (-) (red) or Increase Toxicity / Activity (+) (green), useful for hypotheses and mechanistic interpretations (OECD Principle 5).

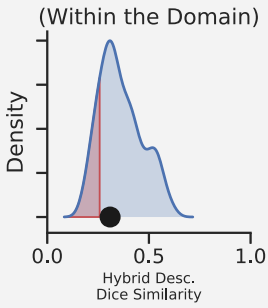
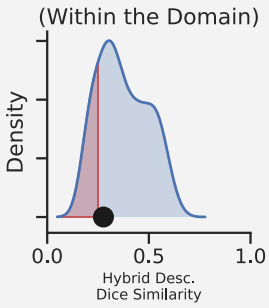
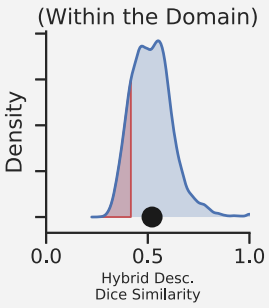
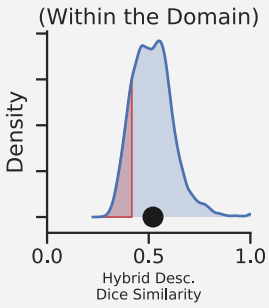
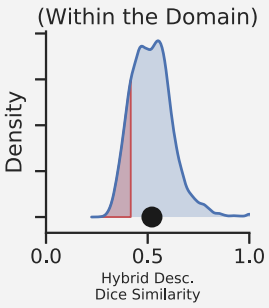
Assay/Method	Prediction class (Confidence)	STR Contribution Mapping
***in vivo* Development Toxicity (Mammals)** Deep Learning categorical model implemented with hybrid descriptors.	Toxic (+) 100.0%	
***in vivo* Reproductive Toxicity (Mammals)** Deep Learning categorical model implemented with hybrid descriptors.	Non-Toxic (-) 100.0%	
***in vitro* Agonist Activity - Estrogen Receptor (ER)** Deep Learning categorical model implemented with hybrid descriptors.	Inactive (-) 71.8%	
***in vitro* Antagonist Activity - Estrogen Receptor (ER)** Deep Learning categorical model implemented with hybrid descriptors.	Inactive (-) 75.4%	
***in vitro* Binding Activity - Estrogen Receptor (ER)** Deep Learning categorical model implemented with hybrid descriptors.	Inactive (-) 83.9%	

Visual AD Inspection®

The applicability domain (AD) is defined by the chemical structure space and the toxicological response encoded by the developed model, to make new predictions with a given reliability (a defined domain of applicability - OECD Principle 3).[2] Our visual AS Inspection® is used to establish the scope and limitations of the models. Basically, new chemicals must be reasonably similar to training set compounds or a prediction cannot be accepted.

Our visual AD inspection is represented by a density plot of the average fingerprint-dice similarity for the k-nearest neighbors of each compound during the 5-Fold external validation. The chemical structure is represented by a hybrid descriptor composed by ECFP6 fingerprint and physicochemical measurements: molecular weight (MW), topological polar surface area (TPSA), octanol-water partition coefficient for neutral compounds ($logK_{ow}$) or at different pH states (logD). At the visual AD inspection, the black circle represents the evaluated compound, the highlighted red area represents the restricted similarity region, and the blue region is the allowed similarity of the chemical space to predict new compounds.

Even though a well-designed AD helps the user to assess the reliability of predictions made by the model, it should not automatically be assumed that all predictions within the defined AD are necessarily reliable.

in vivo **Development Toxicity (Mammals)**	***in vivo*** **Reproductive Toxicity (Mammals)**	***in vitro*** **Agonist Activity - Estrogen Receptor (ER)**	***in vitro*** **Antagonist Activity - Estrogen Receptor (ER)**	***in vitro*** **Binding Activity - Estrogen Receptor (ER)**
				

Final Result

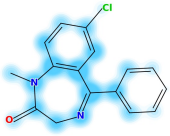
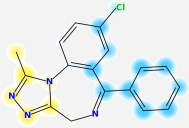
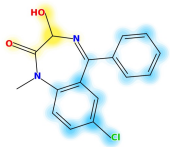
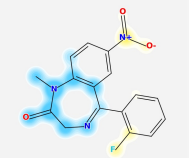
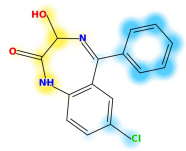
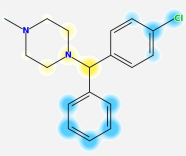
The results are presented below; Toxic (+) and Active (+) (red), and/or Non-Toxic (-) and Inactive (-) (green) predictions are presented with applicability domain (AD) and confidence level (robustness, OECD principle 4) for the Advanced Statistical systems.

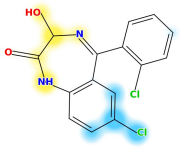
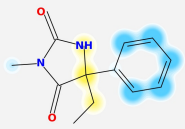
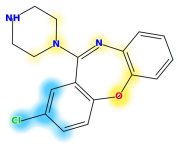
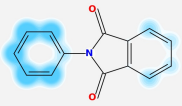
Assay/Method	Result	Applicability Domain
***in vivo* Development Toxicity (Mammals)** Deep Learning categorical model implemented with hybrid descriptors.	**Toxic (+)** **100.0%**	Within 
***in vivo* Reproductive Toxicity (Mammals)** Deep Learning categorical model implemented with hybrid descriptors.	**Non-Toxic (-)** **100.0%**	Within 
***in vitro* Agonist Activity - Estrogen Receptor (ER)** Deep Learning categorical model implemented with hybrid descriptors.	**Inactive (-)** **71.8%**	Within 
***in vitro* Antagonist Activity - Estrogen Receptor (ER)** Deep Learning categorical model implemented with hybrid descriptors.	**Inactive (-)** **75.4%**	Within 
***in vitro* Binding Activity - Estrogen Receptor (ER)** Deep Learning categorical model implemented with hybrid descriptors.	**Inactive (-)** **83.9%**	Within 

Additional Information

Hybrid descriptor Dice similarity was used to improve the deep learning confidence by interpolating the confidence equalized by the compound similarity criteria obtained from the dataset chemical space. This helps to improve the *in silico* toxicological model to reduce the false positive and negative error.

Performance for the 10-most similar molecules

Molecule (Similarity)	Experimental Data	Development Prediction (Confidence)	Reproductive Prediction (Confidence)	Agonist Prediction (Confidence)	Antagonist Prediction (Confidence)	Binding Prediction (Confidence)
 (1.0)	Toxic (+)	Toxic (+) 100.0%	Non-Toxic (-) 100.0%	Inactive (-) 71.78%	Inactive (-) 75.45%	Inactive (-) 83.94%
 (0.58)	Toxic (+)	Toxic (+) 99.97%	Non-Toxic (-) 99.95%	Inactive (-) 79.96%	Inactive (-) 88.73%	Inactive (-) 76.13%
 (0.58)	Toxic (+)	Toxic (+) 99.99%	Non-Toxic (-) 99.88%	Inactive (-) 56.38%	Inactive (-) 58.5%	Inactive (-) 77.42%
 (0.5)	Toxic (+)	Toxic (+) 99.78%	Non-Toxic (-) 99.96%	Inactive (-) 83.78%	Inactive (-) 70.71%	Inactive (-) 88.34%
 (0.43)	Non-Toxic (-)	Toxic (+) 99.83%	Non-Toxic (-) 100.0%	Inactive (-) 63.51%	Inactive (-) 55.83%	Inactive (-) 62.35%
 (0.33)	Toxic (+)	Toxic (+) 99.85%	Toxic (+) 100.0%	Inactive (-) 69.44%	Inactive (-) 79.16%	Inactive (-) 81.93%

Molecule (Similarity)	Experimental Data	Development Prediction (Confidence)	Reproductive Prediction (Confidence)	Agonist Prediction (Confidence)	Antagonist Prediction (Confidence)	Binding Prediction (Confidence)
 (0.32)	Toxic (+)	Toxic (+) 98.54%	Non-Toxic (-) 100.0%	Inactive (-) 83.18%	Inactive (-) 73.76%	Inactive (-) 71.38%
 (0.31)	Toxic (+)	Toxic (+) 100.0%	Toxic (+) 96.4%	Inactive (-) 64.58%	Inactive (-) 82.82%	Inactive (-) 83.04%
 (0.3)	Toxic (+)	Toxic (+) 99.99%	Toxic (+) 100.0%	Inactive (-) 73.95%	Inactive (-) 87.29%	Inactive (-) 83.13%
 (0.3)	Toxic (+)	Non-Toxic (-) 65.12%	Toxic (+) 95.4%	Active (+) 53.71%	Inactive (-) 56.47%	Active (+) 60.88%