	QMRF identifier (JRC Inventory): **Q17-42-0028**
	QMRF Title: **QSAR for acute oral toxicity (in vitro)**
	Printing Date: **Apr 16, 2018**

1.QSAR identifier

1.1.QSAR identifier (title):

QSAR for acute oral toxicity (in vitro)

1.2.Other related models:

1.3.Software coding the model:

QSARModel 3.5.0

Molcode Ltd., Turu 2, Tartu, 51014, Estonia

http://www.molcode.com

2.General information

2.1.Date of QMRF:

30.08.2009

2.2.QMRF author(s) and contact details:

Molcode model development team Molcode Ltd. Turu 2, Tartu, 51014, Estonia

models@molcode.com http://www.molcode.com

2.3.Date of QMRF update(s):

2.4.QMRF update(s):

2.5.Model developer(s) and contact details:

Molcode model development team Molcode Ltd. Turu 2, Tartu, 51014, Estonia

models@molcode.com http://www.molcode.com

2.6.Date of model development and/or publication:

30.08.2009

2.7.Reference(s) to main scientific papers and/or software package:

[1]Karelson M, Dobchev D, Tamm T, Tulp I, Jänes J, Tämm K, Lomaka A, Savchenko D & Karelson G (2008). Correlation of blood-brain penetration and human serum albumin binding with theoretical descriptors. ARKIVOC 16, 38-60.

[2]Karelson M, Karelson G, Tamm T, Tulp I, Jänes J, Tämm K, Lomaka A, Savchenko D & Dobchev D (2009). QSAR study of pharmacological permeabilities. ARKIVOC 2, 218–238.

2.8.Availability of information about the model:

Model is proprietary, but the training and test sets are available.

2.9.Availability of another QMRF for exactly the same model:

None to date

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Rat and mouse

3.2.Endpoint:

[1]QMRF 4. 2. Acute Oral toxicity OECD 420 Acute Oral Toxicity-Fixed Dose Method

[2]QMRF 4. 2. Acute Oral toxicity OECD 423 Acute Oral Toxicity - Acute Toxic Class Method

3.3.Comment on endpoint:

3.4.Endpoint units:

3.5.Dependent variable:

log(IC50) - logarithm of the half maximal inhibitory concentration (IC50) in vitro. The IC50 indicates how much of a particular substance (inhibitor) is needed to inhibit a given biological process (or component of a process) by half.

3.6.Experimental protocol:

Acute oral toxicity is determined using the OECD 420 and OECD 423 (EU B.1.bis. and 1.B.tris.) test guidelines. Acute oral toxicity testing allows to obtain the information on the biologic/toxic activity of a chemical. Currently, the basis for toxicologic classification of chemicals is the median lethal dose (LD50, mg/kg b.w.), which is defined as the statistically derived dose required to kill half the members of a tested population. Animals are observed individually after dosing at least once during the first 30 minutes, periodically during the first 24 hours, with special attention given during the first 4 hours, and daily thereafter, for a total of 14 days.

3.7.Endpoint data quality and variability:

The data were taken from Freidig et al (2001). There are 50 compounds in the dataset of the in vitro cytotoxicity (median of several IC50 values of different experiments) consisting of drugs, agrochemicals and industrial chemicals. The IC50 values of tested substances were translated to logarithmic scale (logIC50) to reduce the range of the data

Statistics:
max value: 2.97
min value: -4.27
standard deviation: 1.90
skewness: -1.00

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSAR

4.2.Explicit algorithm:

Multilinear regression QSAR
Log(IC50) = 5.48 -41.64 * Global softness: 1/(LUMO - HOMO) (AM1)-1.80 * Number of carbonyl groups - 0.30 * Kier&Hall index (order 2)

4.3.Descriptors in the model:

[1]Global softness: 1/(LUMO - HOMO) (AM1)
[2]Number of carbonyl groups
[3]Kier&Hall index (order 2)

4.4.Descriptor selection:

Initial pool of ~1000 descriptors. Stepwise descriptor selection based on a set of statistical selection rules:
1-parameter equations: Fisher criterion and R2 over threshold, variance and t-test value over threshold, intercorrelation with another

descriptor not over threshold
2 parameter equations: intercorrelation coefficient bellow threshold, significant correlation with endpoint in terms of correlation coefficient and t-test.

Stepwise trial of additional descriptors not significantly correlated to any already in the model.

4.5.Algorithm and descriptor generation:

1D, 2D, and 3D theoretical calculations. Quantum chemical descriptors derived from AM1 calculation. Model developed by using multilinear regression.

4.6.Software name and version for descriptor generation:

QSARModel 3.5.0
Molcode Ltd., Turu 2, Tartu, 51014, Estonia
http://www.molcode.com

4.7.Chemicals/Descriptors ratio:

15 (45 chemicals / 3 descriptors)

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

Applicability domain based on training set:
a) by chemical identity: diverse set of chemically reactive organic compounds (alcohols, carboxylic acids, nitriles, aromatic compounds, sulfur and phosphorous compounds, etc)
b) by descriptor value range: This model is suitable for compounds that have the descriptors in the following range:
Global softness: 1/(LUMO - HOMO) (AM1)(min: 0.047 , max: 0.169)
Number of carbonyl groups (min: 0 , max: 1)
Kier&Hall index (order 2) (min: 0 , max: 24.362)

5.2.Method used to assess the applicability domain:

Presence of functional groups in structures
Range of descriptor values in training set with ±30% confidence
Descriptor values must fall between maximal and minimal descriptor values of training set ± 30%

5.3.Software name and version for applicability domain assessment:

QSARModel 3.5.0
Molcode Ltd., Turu 2, Tartu, 51014, Estonia
http://www.molcode.com

5.4.Limits of applicability:

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

Yes

6.2.Available information for the training set:

CAS RN: Yes
Chemical Name: Yes

Smiles: No

Formula: No

INChI: No

MOL file: Yes

6.3.Data for each descriptor variable for the training set:

All

6.4.Data for the dependent variable for the training set:

All

6.5.Other information about the training set:

45 data points: 14 negative: 31 positive values

6.6.Pre-processing of data before modelling:

6.7.Statistics for goodness-of-fit:

R^2= 0.85 (Correlation coefficient);

s = 0.59 (Standard error of the estimate);

F = 76.57 (Fisher statistic);

6.8.Robustness - Statistics obtained by leave-one-out cross-validation:

R^2cv = 0.78 LOO;

6.9.Robustness - Statistics obtained by leave-many-out cross-validation:

R^2cv = 0.83 LMO;

6.10.Robustness - Statistics obtained by Y-scrambling:

6.11.Robustness - Statistics obtained by bootstrap:

6.12.Robustness - Statistics obtained by other methods:

ABC analysis (2:1 training : prediction) on sorted (in increasing order of endpoint value) data divided into 3 subsets (A;B;C). Training set formed with 2/3 of the compounds (set A+B, A+C, B+C) and validation set consisted of 1/3 of the compounds (C, B, A)

Average R^2(fitting) = 0.862

Average R^2(prediction) = 0.802

7.External validation - OECD Principle 4

7.1.Availability of the external validation set:

Yes

7.2.Available information for the external validation set:

CAS RN: Yes

Chemical Name: Yes

Smiles: No

Formula: No

INChI: No

MOL file: Yes

7.3.Data for each descriptor variable for the external validation set:

All

7.4.Data for the dependent variable for the external validation set:

All

7.5.Other information about the external validation set:

5 data points: 2 negative values: 3 positive values

7.6.Experimental design of test set:

The full experimental dataset was sorted according to increasing values of logIC50 and each tenth compound was assigned to the test set.

7.7.Predictivity - Statistics obtained by external validation:

R^2=0.802

7.8.Predictivity - Assessment of the external validation set:

The descriptors for the test set are in the limit of applicability

7.9.Comments on the external validation of the model:

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

The acute oral toxicity depends on the stability and reactivity of the compound, the number of carbonyl groups and the shape of the molecule. Toxicity increases with increasing values of the descriptor Global softness: 1/(LUMO-HOMO)(AM1). The presence of the carbonyl group in the molecule accounts for a higher toxicity. The descriptor Kier&Hall index (order 2) shows that toxicity is also influenced by the shape and size of the molecule.

8.2.A priori or a posteriori mechanistic interpretation:

A posteriori mechanistic interpretation, consistent with published scientific interpretations of experiments.

8.3.Other information about the mechanistic interpretation:

The descriptor Global softness: 1/(LUMO - HOMO) (AM1)[1/eV] gives information about the reactivity and stability of the molecule. Increasing values for Global softness indicates a higher reactivity, so toxicity increases for more reactive compounds. The descriptor Number of carbonyl groups accounts for the presence of carbonyl groups in the molecule. The Kier&Hall index (order 2) gives information about different aspects of atom connectivity within a molecule, about the branching of the ring structures and the flexibility. The proposed mechanism based on the model agrees well with literature (Freidig et al, 2007).

9.Miscellaneous information

9.1.Comments:

9.2.Bibliography:

[1]Freidig AP, Dekkers S, Verwei M, Zvinavashe E, Bessems JGM & van de Sandt JJM (2007). Development of a QSAR for worst case estimates of acute toxicity of chemically reactive compounds. Toxicology Letters 170, 214-222.

[2]NIEHS (2001). Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity Based on Recommendations from an International Workshop Organized by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), NIH Publication No. 01-4500.

9.3.Supporting information:

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

Q17-42-0028

10.2.Publication date:

2017-09-21

10.3.Keywords:

Molcode;acute oral toxicity;cytotoxicity;

10.4.Comments:

former Q8-10-14-176