	QMRF identifier (JRC Inventory): **Q13-412-0066**
	QMRF Title: **Toxtree QSAR 8: rodent carcinogenicity of aromatic amines**
	Printing Date: **Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

Toxtree QSAR 8: rodent carcinogenicity of aromatic amines

1.2.Other related models:

1.3.Software coding the model:

Toxtree

Standalone software application downloadable from the Joint Research Centre (JRC) website http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE

2.General information

2.1.Date of QMRF:

June 2010

2.2.QMRF author(s) and contact details:

[1]Romualdo Benigni Experimental and Computational Carcinogenesis Unit, Environment and Health Department, Istituto Superiore di Sanita Viale Regina Elena 299, 00161 Rome, Italy rbenigni@iss.it

[2]Cecilia Bossa Experimental and Computational Carcinogenesis Unit, Environment and Health Department, Istituto Superiore di Sanita Viale Regina Elena 299, 00161 Rome, Italy cecilia.bossa@iss.it

2.3.Date of QMRF update(s):

2.4.QMRF update(s):

2.5.Model developer(s) and contact details:

[1]Romualdo Benigni rbenigni@iss.it

[2]Rainer Franke

2.6.Date of model development and/or publication:

2001; external validation 2006.

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

[1]Franke R, Gruska A, Giuliani A & Benigni R (2001). Prediction of rodent carcinogenicity of aromatic amines: a quantitative structure-activity relationships model. Carcinogenesis 22, 1561-1571.

[2]Benigni R, Bossa C, Netzeva TI & Worth AP (2007). Collection and evaluation of (Q)SAR models for mutagenicity and carcinogenicity. JRC report EUR 22772 EN. Luxembourg, Office for the Official Publications of the European Communities.
http://ecb.jrc.ec.europa.eu/documents/QSAR/EUR_22772_EN.pdf

2.8.Availability of information about the model:

The model is non-proprietary; 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:

Rodents (rats and mice)

3.2.Endpoint:

4.Human Health Effects 4.12.Carcinogenicity

3.3.Comment on endpoint:

Noncarcinogen / Carcinogen (overall negative/positive score from four experimental groups: rat, mice, male, female). A chemical was considered to be a carcinogen if at least one experimental group gave a positive result.

3.4.Endpoint units:

Noncarcinogen / Carcinogen

3.5.Dependent variable:

Carcinogenicity

3.6.Experimental protocol:

Training and test set data retrieved from literature (different laboratories). A significant proportion of the chemicals was tested with the protocol of the US National Toxicology Program

3.7.Endpoint data quality and variability:

The literature sources were critically reviewed by the QMRF authors.

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSAR

4.2.Explicit algorithm:

QSAR

QSAR8

4.3.Descriptors in the model:

[1]L(R) (length) and B5(R) (maximal width) Sterimol parameters (tabulated in (Verloop 1987))

[2]Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular orbital (LUMO) eV PM3 molecular orbital energies

[3]MR3 , MR5 , MR6 Molar Refractivity contributions of substituents in position 3, 5, and 6 to the amino group

[4]I(An), I(NO2), and I(BiBr) Indicator variables that take value = 1 for anilines, for the presence of a NO2 group, and for biphenyls with a bridge between the phenyl rings, respectively.

4.4.Descriptor selection:

The descriptors were selected by starting from a limited range, chosen on a mechanistic basis, and then expanding the list according to the results of the statistical analyses. In addition to the descriptors accepted in the equation, hydrophobicity (logP) was screened as well, but it was not statistically significant.

4.5.Algorithm and descriptor generation:

4.6.Software name and version for descriptor generation:

Daylight

Partial MR calculation, freely available on-line in 2001 – 2006.

Daylight

http://www.daylight.com/

4.7.Chemicals/Descriptors ratio:

16 = 64 chemicals / 4 descriptors

(11 originally screened)

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

The applicability domain of the model is defined in terms of:

a) structures to which it applies;

b) range of values of the descriptors in the model.

In the comparison of the test to the training set, mathematical chemical structural similarity was also applied.

5.2.Method used to assess the applicability domain:

Structures were checked by human experts. Ranges of descriptors values were calculated. Mathematical chemical similarity between test and training set was assessed as follows. The training and test sets were combined, and the overall Tanimoto similarity matrix was calculated with the Leadscope software. A Euclidian distance matrix was calculated from the similarity matrix and then subjected to Principal Component Analysis. Finally, the ranges of PC scores for the training and test sets were compared.

5.3.Software name and version for applicability domain assessment:

Leadscope v. 2.4

Used for the calculation of the Tanimoto chemical similarity index.

http://www.leadscope.com/

5.4.Limits of applicability:

The types of structures to which the model applies are detailed in the original publication Franke et al.(2001), ref.1 section 9.2; anilines, biphenyl, naphthalenes, fluorenes.

Ranges of descriptors for the training set:

L(R) between 2.060 and 5.970

B5(R) between1.000 and 4.040

HOMO between -9.544 and -7.990

LUMO between -1.594 and 0.438

MR3 between 0.100 and 0.800

MR5 between 0.090 and 1.490

MR6 between 0.090 and 0.600

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: Yes

Formula: Yes

INChI: No

MOL file: No

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:

64 data points: 52 positive values; 12 negative values

6.6.Pre-processing of data before modelling:

6.7.Statistics for goodness-of-fit:

Squared Canonical Correlation = 0.50. The equation correctly reclassified 93.7.% (Accuracy) of the compounds (Class1, noncarcinogens, 92.7% (Specificity); Class2, carcinogens, 94.2% (Sensitivity)).

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

Same as above

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

Three leave-many-out procedures were applied, leaving out:

a) 10%;

b) 25%; and

c) 50% of the training set.

In addition, each procedure was applied in two different ways, by generating test sets with the following characteristics:

1) Option 1: with the same proportion Class1/Class2 present in the whole sample of chemicals;

2) Option 2: without the above constraint. Each procedure was applied ten times.

Option 1:

Sensitivity (10%: 80.0; 25%: 84.6; 50%: 83.8)

Specificity (10%: 70.0; 25%: 80.0; 50%: 81.7)

Accuracy (10%: 78.3; 25%: 83.8; 50%: 83.4)

Option 2:Sensitivity (10%: 85.2; 25%: 79.9; 50%: 82.5)

Specificity (10%: 71.7; 25%: 76.2; 50%: 73.1)Accuracy (10%: 81.7; 25%: 78.1; 50%: 80.3)

6.10.Robustness - Statistics obtained by Y-scrambling:

6.11.Robustness - Statistics obtained by bootstrap:

6.12.Robustness - Statistics obtained by other methods:

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: Yes

Formula: Yes

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:

External validation set with 27 compounds appended: 13 noncarcinogens, 14 carcinogens

7.6.Experimental design of test set:

The test set consisted of literature data retrieved after the original model was published.

The applicability domain was taken into account by considering chemical structure, range of descriptors values, chemical similarity.

Ranges of descriptors for the test set:

L(R) between 2.060 and 10.790

B5(R) between 1.000 and 7.990

HOMO between -9.416 and -8.043

LUMO between -1.304 and 0.461

MR3 between 0.100 and 0.800

MR5 between 0.100 and 0.740

MR6 between 0.100 and 0.600

7.7.Predictivity - Statistics obtained by external validation:

Accuracy = 0.70; Sensitivity = 0.92; Specificity = 0.46.

7.8.Predictivity - Assessment of the external validation set:

The test set is representative of the applicability domain; no further chemicals were available from the literature.

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:

Whereas the principal factor that affects the relative carcinogenicity (potency) of the aminoarenes is their hydrophobicity (logP), followed by electronic factors (HOMO and LUMO) and then steric factors, the model for the yes/no activity shows no influence of logP. This indicates that the potential to be active depends on a threshold of reactivity (HOMO and LUMO), and on the steric hindrance at substitution positions 3, 5, and 6 of the ring, together with steric hindrance due to bulky substituents to the nitrogen. The parameters in the model are mechanistically linked to metabolic activation.

8.2.A priori or a posteriori mechanistic interpretation:

The descriptors to be screened were selected based on knowledge of the mechanisms of action. The descriptors actually accepted in the model

confirm the a priori hypotheses.

8.3.Other information about the mechanistic interpretation:

9.Miscellaneous information

9.1.Comments:

9.2.Bibliography:

[1]Franke R, Gruska A,Giuliani A & Benigni R (2001). Prediction of rodent carcinogenicity of aromatic amines: a quantitative structure-activity relationships model. Carcinogenesis 22, 1561-1571.

[2]Benigni R, Bossa C, Jeliazkova N, Netzeva TI & Worth AP (2008). The Benigni / Bossa rulebase for mutagenicity and carcinogenicity – a module of ToxTree. JRC report EUR 23241 EN. Luxembourg: Office for Official Publications of the European Communities,1- 75. http://ecb.jrc.ec.europa.eu/documents/QSAR/EUR_23241_EN.pdf

[3]Beland FA & Kadlubar FF (1990). Metabolic activation and DNA adducts of aromatic amines and nitroaromatic hydrocarbons. In Chemical carcinogenesis and mutagenesis I. Eds. Cooper CS, Grover PL. pp. 265-325. Berlin: Springer Verlag.

[4]Benigni R (2005). Structure-activity relationship studies of chemical mutagens and carcinogens: mechanistic investigations and prediction approaches. Chemical Reviews 105, 1767-1800.

[5]Benigni R, Giuliani A, Gruska A & Franke R (2003). QSARs for the mutagenicity and carcinogenicity of the aromatic amines. In Quantitative Structure-Activivity Relationship (QSAR) models of mutagens and carcinogens. Ed. Benigni R. pp 125-144. Boca Raton: CRC Press.

[6]Debnath AK, Debnath G, Shusterman AJ & Hansch C (1992). A QSAR investigation of the role of hydrophobicity in regulating mutagenicity in the Ames test: 1. Mutagenicity of aromatic and heteroaromatic amines in Salmonella typhimurium TA98 and TA100. Environmental and Molecular Mutagenesis 19, 37-52.

[7]Kadlubar FF & Beland FA (1985). Chemical properties of ultimate carcinogenic metabolites of arylamines and arylamides. In Polycyclic hydrocarbons and carcinogenesis. Ed. Harvey RG. pp 341-370. Washington DC: American Chemical Society.

[8]Kadlubar FF & Hammons GJ (1987). The role of ctochrome P-450 in the metabolism of chemical carcinogens. In Mammalian cytochromes P-450, vol. II. Ed. Guengerich FP. pp 81-130. Boca Raton: CRC Press.

[9]Woo YT & Lai DY (2001). Aromatic amino and nitro-amino compounds and their halogenated derivatives. In Patty's Toxicology. Ed. Bingham E, Cohrssen B, Powell CH. pp 969-1105. New York: John Wiley and Sons.

9.3.Supporting information:

Toxtree QSAR8 training_64.sdf	http://qsardb.jrc.ec.europa.eu/qmrf/protocol/Q13-412-0066/attachment/A764
Toxtree QSAR8 test_27.sdf	http://qsardb.jrc.ec.europa.eu/qmrf/protocol/Q13-412-0066/attachment/A765

Test set(s)

10.Summary (JRC QSAR Model Database)

Supporting information

10.1.QMRF number:

Q13-412-0066

10.2.Publication date:

2013-07-02

10.3.Keywords:

Toxtree;carcinogenicity;aromatic amine;rodent;rat;mouse;

10.4.Comments:

former Q19-35-35-291