	QMRF identifier (JRC Inventory): **Q15-35-0006**
	QMRF Title: **Quantitative Structure Activity Relationship for CHemical Ecotoxicity (QSARCHE) - chronic fish toxicity by polar narcosis**
	Printing Date: **Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

Quantitative Structure Activity Relationship for CHemical Ecotoxicity (QSARCHE) - chronic fish toxicity by polar narcosis

1.2.Other related models:

None

1.3.Software coding the model:

QSARCHE

http://www.arche-consulting.be/organics-toolbox/qsarche-model/

2.General information

2.1.Date of QMRF:

27-03-2012

2.2.QMRF author(s) and contact details:

[1]Lieve Claeys Arche, Stapelplein 70, box 104, 9000 Gent, Belgium lieve.claeys@arche-consulting.be

[2]Federica Iaccino Arche, Stapelplein 70, box 104, 9000 Gent, Belgium federica.iaccino@arche-consulting.be

[3]Frederik Verdonck Arche, Stapelplein 70, box 104, 9000 Gent, Belgium frederik.verdonck@arche-consulting.be

2.3.Date of QMRF update(s):

2.4.QMRF update(s):

2.5.Model developer(s) and contact details:

[1]Lieve Claeys Arche, Stapelplein 70, box 104, 9000 Gent, Belgium lieve.claeys@arche-consulting.be

[2]Frederik Verdonck Arche, Stapelplein 70, box 104, 9000 Gent, Belgium frederik.verdonck@arche-consulting.be

[3]Federica Iaccino Arche, Stapelplein 70, box 104, 9000 Gent, Belgium federica.iaccino@arche-consulting.be

2.6.Date of model development and/or publication:

Model published in 2012

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

2.8.Availability of information about the model:

The model is non-proprietary and freely available. Information on the algorithm, training and test set is publicly 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:

Fish (several freshwater species)

3.2.Endpoint:

3.Ecotoxic effects 3.5.Long-term toxicity to fish (egg/sac fry, growth inhibition of juvenile fish, early life stage, full life cycle)

3.3.Comment on endpoint:

NOEC (No Observed Effect Concentration), measured after minimum 4 days exposure.
Chronic toxicity is required under the REACH regulation. REACH regulation states that a substance is identified as toxic (T) when long term NOEC =< 0.01 or 0.1 mg/L. Thus the endpoint could also be treated in classification. Further thresholds apply for the CLP regulation, and for the chemical safety report (CSR), required by REACH.

3.4.Endpoint units:

NOEC unit is log(mmol/L).

3.5.Dependent variable:

Log NOEC

3.6.Experimental protocol:

Toxicity data were extracted from the ECHA dissemination website (ref 3, section 9.2) and the ECETOC database (ref 4, section 9.2) for the training sets. The test set data come from the ECOTOX, OASIS databases.
Several test protocols are included in the experimental databases.
Experimental data are selected based on
-freshwater species
-duration of test more than 4 days
-mortality, growth and reproduction
-no static tests

3.7.Endpoint data quality and variability:

Data from ECHA dissemination website are of very high quality (Klimisch 1 and 2 score). The ECETOC database includes experimental results for aquatic toxicity from European Centre for Ecotoxicology of Chemicals (ECETOC). The principal quality criteria for acceptance of data were that test methods should be well described and the toxicant concentrations must be measured. Therefore ECETOC provides a large comprehensive compilation of highly reliable ecotoxicity data.
Data extracted from the US EPA ECOTOX (ref 6, section 9.2) and OASIS (ref 5, section 9.2) database are likely to be of variable quality and therefore only used as validation test set.

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSAR

4.2.Explicit algorithm:

Multiple Linear Regression (MLR) based on Log P and pKa descriptors
Log NOEC (mmol/L)= -0.984 – 0.627 Log P + 0.0514 pKa
Polar narcosis: Log NOEC (mmol/L)= -0.984 – 0.627 Log P + 0.0514 pKa

4.3.Descriptors in the model:

[1]Log P dimensionless logarithm of octanol-water partition coefficient and represents the ratio of the solubility of a compound in octanol (a non-polar solvent) to its solubility in water (a polar solvent). Log P data are experimental values.

[2]pKa dimensionless quantitative measure of the strength of an acid in solution. pKa descriptor is based on computed values (ChemAxon via the OECD QSAR toolbox).

4.4.Descriptor selection:

A large set of 2D descriptors (269) was computed using the DRAGON software version 6.0. The 2D and 3D (27) descriptors, which are pre-calculated by the OECD QSAR application toolbox were also included in the pool to select the relevant descriptors.

By combining existing mechanistic understanding and multivariate statistics analysis (i.e. PCA), the key descriptors were selected.

4.5.Algorithm and descriptor generation:

2D descriptors were used. The descriptors data are both experimental (Log P) and calculated values (pKa).

4.6.Software name and version for descriptor generation:

OECD QSAR toolbox version 2.2

includes 2D and 3D parameters calculation. Log P are experimental values, pKa is calculated in the OECD QSAR toolbox, by ChemAxon.

http://www.oecd.org/document/54/0,3746,en_2649_34379_42923638_1_1_1_1,00.html#Download_qsar_application_toolbox

4.7.Chemicals/Descriptors ratio:

polar narcosis: 11 substances/2 descriptors = 5.5

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

Polar narcosis: the QSAR model is defined to be applicable to polar narcosis chemicals (Class II – Verhaar modified, 2011), presenting a Log P range of 1.46 to 5.76 and pKa range of 6.83 to 10.7.

5.2.Method used to assess the applicability domain:

Information for the user: the users have two different ways to evaluate the applicability domain of the model provided by QSARCHE:

1) The user should check manually whether the descriptors of the substance are within the range of the training set descriptors

2) The user should check for similar compounds in the training set which is publically available.

5.3.Software name and version for applicability domain assessment:

QSARCHE

http://www.arche-consulting.be/organics-toolbox/qsarche-model/

5.4.Limits of applicability:

As anticipated in section 5.1:

- Polar narcosis mechanism of chronic fish toxicity (Class 2 Verhaar modified) (ref 2, section 9.2)

It is not possible to process inorganic compounds, mixtures (in addition consider that stereoisomers are not distinguished) and metal complexes.

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: No

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:

The full training set is provided in supporting information.

6.6.Pre-processing of data before modelling:

6.7.Statistics for goodness-of-fit:

$R^2$0.808

R^2adj 0.802

Sum of squared residuals 18.5

Sample standard deviation of residuals 0.510

Fisher function 149

Fisher threshold for statistical significance 3.89

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

Q^2(Leave one out) 0.75

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

Q^2(Leave many out) 0.70

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: No

INChI: No

MOL file: No

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:

The test set is provided in supporting information

7.6.Experimental design of test set:

7.7.Predictivity - Statistics obtained by external validation:

$R^2$0.867

R^2adj 0.861

Q^2(Leave one out) 0.838

Sum of squared residuals 4.69

Sample standard deviation of residuals 0.330

Fisher function 141

Fisher threshold for statistical significance 4.02

7.8.Predictivity - Assessment of the external validation set:

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 model largely relies on logP, which is typically the main descriptor used for non polar narcosis to fish: log P is linked to the partitioning of the substance from water into the lipid phase of the fish organism, thus reflecting the chemical bioavailability. pKa was selected as extra relevant descriptor to Log P for polar narcosis. There are several causes of the influence of the pKa on the toxicity according to Könemann et al (1981)(ref 7, section 9.2). The first one is the influence on the accumulation: the molecular form of an acid can pass through membranes (the gills) faster than the ionic form, in both directions. The primary toxic action of the phenolic compounds will be caused by either the molecular form, or the ionic form. At a given total concentration inside the fish, the concentration of the active form is pKa dependent. Finally, the interaction with the receptor of the active form of the chemical can depend on its electronic situation, which also governs the pKa. In this way pKa and this interaction can be correlated. In contrast to the first one, these last two pKa influences will not depend directly on the external pH.

8.2.A priori or a posteriori mechanistic interpretation:

Based on PCA, literature search and expert mechanistic understanding, relevant descriptors were selected and the mechanistic understanding was confirmed and expanded.

8.3.Other information about the mechanistic interpretation:

9.Miscellaneous information

9.1.Comments:

9.2.Bibliography:

[1]Claeys L, Iaccino F, Van Sprang P, Verdonck F. 2012. Development and validation of a QSAR for chronic narcosis to fish. http://www.ncbi.nlm.nih.gov/pubmed/23775559

[2]Enoch S J, Hewitt M, Cronin M T D, Azam S & Madden J C (2008) Classification of chemicals according to mechanism of aquatic toxicity: an evaluation of the implementation of the Verhaar scheme in Toxtree. Chemosphere 73 (3) 243-248 http://www.mendeley.com/research/classification-chemicals-according-mechanism-aquatic-toxicity-evaluation-implementation-verhaar-scheme-toxtree/

[3]ECHA dissemination website http://echa.europa.eu/web/guest/information-on-chemicals/registered-substances

[4]ECETOC database http://www.ecetoc.org/index.php?mact=MCSoap,cntnt01,details,0&cntnt01by_category=5&cntnt01template=display_list_v2&cntnt01order_by=Number%20Desc&cntnt01display_template=display_details_v2&cntnt01document_id=234&cntnt01returnid=89

[5]OASIS database http://oasis-lmc.org/

[6]ECOTOX database http://cfpub.epa.gov/ecotox/

[7]Könemann H & Musch A (1981) Quantitative structure-activity relationships in fish toxicity studies Part 2: The influence of pH on the QSAR of chlorophenols.Toxicology 19 (3), 223 - 8 http://dx.doi.org/10.1016/0300-483X(81)90131-1

9.3.Supporting information:

Training set(s)Test set(s)Supporting information

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

Q15-35-0006

10.2.Publication date:

2015-03-05

10.3.Keywords:

Chronic fish toxicity;polar narcosis;QSARCHE;

10.4.Comments:

old # Q33-49-44-427