	QMRF identifier (JRC Inventory): **Q15-35-0009**
	QMRF Title: **Chronic fish toxicity model for predicting sub-lethal NOEC values for non-polar narcotics.**
	Printing Date: **Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

Chronic fish toxicity model for predicting sub-lethal NOEC values for non-polar narcotics.

1.2.Other related models:

1.3.Software coding the model:

None

2.General information

2.1.Date of QMRF:

19/5/2014

2.2.QMRF author(s) and contact details:

Tom Austin Shell Health Rsik Science Team +44 (0) 1614994733 tom.austin@shell.com

2.3.Date of QMRF update(s):

2.4.QMRF update(s):

2.5.Model developer(s) and contact details:

[1]Tom Austin Shell Health Risk Science Team +44 (0) 1614994733 tom.austin@shell.com

[2]Charles Eadsforth Shell Health Risk Science Team Charles.eadsforth@shell.com

2.6.Date of model development and/or publication:

17 March 2014 (publication)

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

Austin TJ & Eadsforth CV (2014) Development of a chronic fish toxicity model for predicting sub-lethal NOEC values for non-polar narcotics.SAR and QSAR in Environmental Research.25(2), 147-160 http://dx.doi.org/10.1080/1062936X.2013.871577

2.8.Availability of information about the model:

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

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Freshwater fish included in dataset:

Pimephales promelas

Oryzias latipes

Danio rerio

Oncorhynchus kisutch

Oncorhynchus mykiss

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:

No mortality data were included in the dataset, only sub-lethal endpoints generated using chronic study methodologies suitable for REACH

endpoint fulfilment.

3.4.Endpoint units:

mg/L

3.5.Dependent variable:

log NOEC in mmol/L

3.6.Experimental protocol:

ECHA dissemination portal and USEPA ECOTOX database were used as data sources

3.7.Endpoint data quality and variability:

Classifying chemicals for chronic toxicity under the CLP regulation (the EU's implementation of the United Nation's Globally Harmonised System for classification and labelling) and amendments hitherto [9,10], is carried out by assessing chronic NOEC or EC10 values. When considering environmental risk, the REACH regulations take a conservative approach [1]. Where there are multiple valid toxicity values for an end point, the most conservative (worst case) is always used; additionally, if there is a lack of data across trophic levels, assessment factors may be applied to compensate for the uncertainty. In this vein, and considering the inherent uncertainty in QSAR predictions, it was decided that a conservative approach would be taken and only NOEC data would be gathered and included into the QSAR training set as opposed to LOEC or MATC values. Best practice in QSAR development dictates that only one value should be included for each compound within the training set [4]. To ensure the most conservative model, only the most sensitive end points were included for each substance (growth/reproduction) and mortality NOECs were filtered out at an early stage.

Chronic fish NOEC data were obtained from the following sources:

• The EPA's ECOTOX database (http://cfpub.epa.gov/ecotox/advanced_query.htm).

• The ECETOC database downloaded through the OECD QSAR Toolbox (http://www.oecd.org/env/ehs/risk-assessment/theoecdqsartoolbox.htm).

• The ECHA dossier dissemination portal (http://echa.europa.eu/information-on-chemicals/registered-substances).

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

4.2.Explicit algorithm:

log NOEC (mmol/L) = 0.711 – 0.914 log Kow

Results will be in mmol/L, anti log and multiplication by the molecular weight are required to obtain a result in mg/L

log NOEC (mmol/L) = 0.711 – 0.914 log Kow

4.3.Descriptors in the model:

Log Kow The partition coefficient between 1-Octanol and water

4.4.Descriptor selection:

Log Kow was used as the only descriptor for the model. Its selection was based on a vast scientific literature demonstrating the relationship between log Kow and the aquatice toxicity of chemicals acting via non-polar narcosis. No other descriptors were deemed necessary based on the R2 value obtained when correlating the toxicity data obtained with log Kow.

4.5.Algorithm and descriptor generation:

Experimental log Kow values for each compound were taken from either their ECHA dossier or the ECOSAR (v1.11) PhysProp database. Data were only taken from the ECHA dossier from studies rated reliable under the Klimisch scoring system (i.e. 1 or 2) and performed using a suitable test method e.g. OECD 123. The log Kow of each compound was plotted (x-axis) against the NOEC value (y-axis, mmol/L). The algorithm was taken from the resulting linear regression.

4.6.Software name and version for descriptor generation:

4.7.Chemicals/Descriptors ratio:

19 chemicals/ 1 descriptor

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

To fall within the applicability domain of this model, compounds must:
exert toxicity via non-polar narcosis,
have a log Kow value between 0.46 and 5.30.

Molecular weights of chemicals included in the training set range from 72.11 to 258.41.

Due to the strong and simple linear relationship, it is anticipated that reasonable predicitons could be made outside this domain, however this is at the users own risk.

5.2.Method used to assess the applicability domain:

See supporting information.

5.3.Software name and version for applicability domain assessment:

5.4.Limits of applicability:

See supporting information.

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:

Full data set available in publication and sdf file

6.6.Pre-processing of data before modelling:

The lowest (most conservative) reliable value was used per compound.

6.7.Statistics for goodness-of-fit:

r^2 0.91

r^2_{adj} 0.90

S(est) 0.38

Fisher statistic 165

p <0.001

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

Q^2_{LOO} 0.88

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

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:

Full data set available in publication and sdf file

7.6.Experimental design of test set:

Model processing and statistical analysis were performed using Minitab 16 Statistical Software (http://www.minitab.com/en-us/products/minitab/). Potential outliers within the preliminary dataset were identified using regression analysis to identify data points with a high normalized residual value (>2). These data points were subject to further investigation to check whether they were indeed outliers. To allow external validation of the model the

data were divided into a training set and a test set. Dearden et al. [4](see attached publication for reference) discuss the pitfalls that can occur in this process and suggest a rational approach be adopted; additionally, a 2:1 training to test set ratio is recommended. Here, the full data set was ordered by log Kow from low to high. Then, going from low to high, every third compound was taken and placed separately for the test set (excluded from model creation). This method allowed the created model to be tested across the full range of log Kow values and ensured the resulting external validation parameters were not unfairly dependent on the test-set selection. The lowest and highest log Kow compounds were left in the training set each time in order to keep the largest applicability domain possible, therefore the last compound selected for the test set was the second to last rather than the last.

7.7.Predictivity - Statistics obtained by external validation:

r^2_{ext} 0.89

S(pre) 0.41

7.8.Predictivity - Assessment of the external validation set:

All chemicals in test set were within the applicability domain and were spread across to allow assessment of predicitons throughout the domain.

7.9.Comments on the external validation of the model:

See supporting information (publication).

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

Log Kow relationship with aquatic toxicity exerted via non-polar narcotics.

See publication attached.

8.2.A priori or a posteriori mechanistic interpretation:

A priori

8.3.Other information about the mechanistic interpretation:

9.Miscellaneous information

9.1.Comments:

9.2.Bibliography:

Austin TJ & Eadsforth CV (2014) Development of a chronic fish toxicity model for predicting sub-lethal NOEC values for non-polar narcotics.SAR and QSAR in Environmental Research.25(2), 147-160 http://dx.doi.org/10.1080/1062936X.2013.871577

9.3.Supporting information:

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

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

Q15-35-0009

10.2.Publication date:

2015-03-05

10.3.Keywords:

Chronic fish toxicity;NOEC;non-polar narcotic;

10.4.Comments:

old # Q30-45-40-421