	QMRF identifier (JRC Inventory): **To be entered by JRC**
	QMRF Title: **Developmental Toxicity model (CAESAR) - v. 2.1.8**
	Printing Date: **7-10-2022**

1.QSAR identifier

1.1.QSAR identifier (title):

Developmental Toxicity model (CAESAR) - v. 2.1.8

1.2.Other related models:

Several models based on classification methods and hybrid techniques were developed within CAESAR project. These models, developed on the same dataset, are able to predict developmental toxicity with good performances.

1.3.Software coding the model:

VEGA (https://www.vegahub.eu/)

The VEGA software provides QSAR models to predict tox, ecotox, environ, phys-chem and toxicokinetic properties of chemical substances.

emilio.benfenati@marionegri.it

2.General information

2.1.Date of QMRF:

7-10-2022

2.2.QMRF author(s) and contact details:

[1] Antonio Cassano Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy antonio.cassano@marionegri.it https://www.marionegri.it/

[2] Emilio Benfenati Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy emilio.benfenati@marionegri.it https://www.marionegri.it/

2.3.Date of QMRF update(s):

NA

2.4.QMRF update(s):

CAESAR Developmental Toxicity Model 2.0 This is the standalone version of the CAESAR Developmental Toxicity Model 1.0. This software implements only the Developmental Toxicity endpoint. The Applicability Domain tab is the main improvement to the previous version.

2.5.Model developer(s) and contact details:

[1] Emilio Benfenati Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy emilio.benfenati@marionegri.it https://www.marionegri.it/

[2] Antonio Cassano Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy antonio.cassano@marionegri.it https://www.marionegri.it/

[3] Alberto Manganaro Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy alberto.manganaro@marionegri.it https://www.marionegri.it/

2.6.Date of model development and/or publication:

The model was published in 2010.

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

[1] Cassano, A., Manganaro, A., Martin, T. et al. CAESAR models for developmental toxicity. Chemistry Central Journal 4, S4 (2010). https://doi.org/10.1186/1752-153X-4-S1-S4

[2] Benfenati E, Roncaglioni A, Lombardo A, Manganaro A. Integrating QSAR, Read-Across, and Screening Tools: The VEGAHUB Platform as an Example. Advances in Computational Toxicology; Springer; 2019. p. 365-81.

2.8.Availability of information about the model:

The model is non-proprietary and the training set is available.

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

Another QMRF is not available.

3.Defining the endpoint - OECD Principle 1

3.1.Species:

.

Human and rat

3.2.Endpoint:

TOX 7.8.2. Developmental toxicity

3.3.Comment on endpoint:

.NA

3.4.Endpoint units:

Adimensional.

3.5.Dependent variable:

CAESAR binary class: Non developmental toxicant / Developmental toxicant

3.6.Experimental protocol:

NA

3.7.Endpoint data quality and variability:

The developmental toxicity data employed in CAESAR project comprises 292 compounds extracted from Arena et al. (see [1] in 9.2) with their names, CAS numbers, molecular structures and toxicity classes. This developmental toxicity database was constructed by combining subsets of information from the Teratogen Information System (TERIS) [2] and US Food and Drug Administration (FDA) guidelines [3]. Both sources are on evaluation of the existing human and animal data on potentially teratogenic chemicals, which are used by physicians for reference. The TERIS compilation is skewed toward a complete evaluation of the animal data; whereas the FDA discussions emphasize human studies or case reports, with some reference to pertinent animal studies.

The original data set includes 293 compounds, but Azatguiorube was eliminated because there was no structural information about this compound in two databases of chemical structures: Chemfinder and ChemIDPlus. Then he dataset was individually checked in order to be sure that the chemical structures to be used for modeling were correct. Finally, we removed inorganic ions and water molecules.

The developmental toxicity data set (see [1] in 9.2) was firstly subdivided into 5 categories, according to the FDA criteria. Then, for developing classification models, the developmental toxicity data set was subdivided in two classes, i.e. non developmental toxicant (N) and developmental toxicant (D). The class N merges the first FDA two categories i.e. Category A and B, whereas the class D includes all compounds belonging to categories C, D and X.

Finally, the data set was split into training (234 substances) and test sets (58 substances) using rational design, by CAESAR Partner Helmholtz-Zentrum für Umweltforschung, using ChemProp [4,5].

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSAR classification model for Developmental Toxicity based on a Random Forest method implemented using WEKA open-source libraries.

4.2.Explicit algorithm:

Random forest

A random forest is a classifier consisting of a collection of tree-structured classifiers such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalization error for forests converges to a limit as the number of trees in the forest becomes large. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them [6].

The Random Forest classifier is implemented as in the WEKA libraries version 3.5.8 using the following parameters: Number of trees in the forest: 10, Random number seed: 1, Depth: 12

4.3.Descriptors in the model:

[1] SdssC adimensional Sum of all (C –) E-State values in molecule

[2] SHssNH Sum of all [– NH –] E-State values in molecule

[3] icycem Mean Information on the Vertex Cycle Matrix y

[4] BEHm1 highest eigenvalue n.1 of Burden matrix / weighted by atomic masses

[5] BELv1 lowest eigenvalue n. 1 of Burden matrix / weighted by atomic van der Waals volumes

[6] BELv8 lowest eigenvalue n.8 of Burden matrix / weighted by atomic van der Waals volumes

[7] BELp3 lowest eigenvalue n.3 of Burden matrix / weighted by atomic polarizabilities

[8] MATS4v Moran autocorrelation - lag4 / weighted by atomic van der Waals volumes

[9] MATS1p Moran autocorrelation - lag1 / weighted by atomic polarizabilities

[10] MATS4p Moran autocorrelation - lag4 / weighted by atomic polarizabilities

[11] GATS2m Geary autocorrelation - lag 2/ weighted by atomic polarizabilities

[12] GATS3v Geary autocorrelation - lag 3 / weighted by atomic van der Waals volumes

[13] GATS1p Geary autocorrelation - lag1 / weighted by atomic polarizabilities

4.4.Descriptor selection:

To avoid false chance correlation between the descriptors and biological activity, a variety of methods were employed in CAESAR project to reduce the data ‘noise’. The technique used for feature selection was based on the multileveled-self organization. The goal of multileveled-self-organization is to select a composition of most relevant input variables with respect to best satisfying the final goal of modelling.

4.5.Algorithm and descriptor generation:

EPA descriptors have been used for modeling. They refer to descriptors calculated using Toxicity Estimation Software Tool (T.E.S.T.). The selected number of descriptors is 13.

4.6.Software name and version for descriptor generation:

Descriptors are calculated by an in-house JAVA software, developed by Todd Martin (EPA), based on the CDK open-source libraries.

4.7.Chemicals/Descriptors ratio:

234/13 = 18

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

The Applicability Domain (AD) is assessed using the original algorithm implemented within VEGA. An overall AD index is calculated, based on a number of parameters, which relate to the results obtained on similar chemicals within the training and test sets.

ADI is defined in this way for this QSAR model´s predictions:

If 1 ≥ AD index ≥ 0.8, the predicted substance is regarded in the Applicability Domain of the model. It corresponds to “good reliability” of prediction.

If 0.8 > AD index ≥ 0.7, the predicted substance could be out of the Applicability Domain of the model. It corresponds to “moderate reliability” of prediction.

If AD index < 0.7, the predicted substance is regarded out of the Applicability Domain of the model. It corresponds to “low reliability” of prediction.

5.2. Method used to assess the applicability domain:

The VEGA applicability domain and chemical similarity are measured with the algorithm developed for VEGA. Full details are in the VEGA website (www.vegahub.eu), including the open access paper describing it [7]. The AD also evaluates the correctness of the prediction on similar compounds (accuracy), the consistency between the predicted value for the target compound and the experimental values of the similar compounds, the range of the descriptors, and the presence of unusual fragments, using atom centred fragments.

These indices are defined in this way for this QSAR model:

Similar molecules with known experimental value:

This index takes into account how similar are the first two most similar compounds found. Values near 1 mean that the predicted compound is well represented in the dataset used to build the model, otherwise the prediction could be an extrapolation. Defined intervals are:

If 1 ≥ index > 0.8, strongly similar compounds with known experimental value in the training set have been found

If 0.8 ≥ index > 0.7, only moderately similar compounds with known experimental value in the training set have been found

If index ≤ 0.7, no similar compounds with known experimental value in the training set have been found

Accuracy (average error) of prediction for similar molecules:

This index takes into account the classification accuracy in prediction for the two most similar compounds found. Values near 1 mean that the predicted compounds fall in an area of the model's space where the model gives reliable predictions (no misclassifications), otherwise the lower is the value, the worse the model behaves. Defined intervals are:

If 1 ≥ index > 0.8, accuracy of prediction for similar molecules found in the training set is good

If 0.8 ≥ index > 0.6, accuracy of prediction for similar molecules found in the training set is not optimal

If index ≤ 0.6, accuracy of prediction for similar molecules found in the training set is not adequate

Concordance for similar molecules:

This index takes into account the difference between the predicted value and the experimental values of the two most similar compounds. Values near 0 mean that the prediction made disagrees with the values found in the model's space, thus the prediction could be unreliable. Defined intervals are:

If 1 ≥ index > 0.8, molecules found in the training set have experimental values that agree with the target compound predicted value

If 0.8 ≥ index > 0.6, similar molecules found in the training set have experimental values that slightly disagree with the target compound predicted value

If index ≤ 0.6, similar molecules found in the training set have experimental values that completely disagree with the target compound predicted value

Atom Centered Fragments similarity check:

This index takes into account the presence of one or more fragments that aren't found in the training set, or that are rare fragments. First order atom centered fragments from all molecules in the training set are calculated, then compared with the first order atom centered fragments from the predicted compound; then the index is calculated as following

If index = 1, all atom centered fragment of the compound have been found in the compounds of the training set

If 1 > index ≥ 0.7, some atom centered fragment of the compound have not been found in the compounds of the training set or are rare fragments

If index < 0.7, a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments

Atom Centered Fragments similarity check:

This index checks if the descriptors calculated for the predicted compound are inside the range of descriptors of the training and test set. The index has value 1 if all descriptors are inside the range, 0 if at least one descriptor is out of the range. Defined intervals are:

If index = True, descriptors for this compound have values inside the descriptor range of the compounds of the training set

If index = False, descriptors for this compound have values outside the descriptor range of the compounds of the training set

5.3.Software name and version for applicability domain assessment:

VEGA (www.vegahub.eu)

5.4.Limits of applicability:

The model is not applicable to inorganic chemicals and substances containing unusual elements (i.e., different from C, O, N, S, P, Cl, Br, F, I). Salts can be predicted only if converted to the neutralized form.

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: Yes

NanoMaterial: 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:

Dataset n = 292

6.6.Pre-processing of data before modelling:

See 3.3

6.7.Statistics for goodness-of-fit:

Training set: n = 234

Accuracy 100%; FP rate 0%; FN rate 0%; PPV 100%; NPV 100%; Sensitivity 100%; Specificity 100%; Nb unpredicted 0.

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

NA

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

10-folds-CV on the training set: Accuracy 77%; FP rate 39%; FN rate 16%; PPV 82%; NPV 64%; Sensitivity 84%; Specificity 61%; Nb unpredicted 54.

6.10.Robustness - Statistics obtained by Y-scrambling:

NA

6.11.Robustness - Statistics obtained by bootstrap:

NA

6.12.Robustness - Statistics obtained by other methods:

NA

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: Yes

NanoMaterial: 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:

NA

7.6.Experimental design of test set:

The dataset was splitted in training and test sets in a rational way. All the compounds were sorted according to a hierarchical system of compound classes in relation to functional group. Within classes, the compounds were sorted according to halogen substitution, aromaticity, bond orders, ring contents, and number of atoms. Particular attention was paid to ordering compounds with mixed functional property groups.

The test set was separated from the sorted list by keeping the relations between these compound classes in both resulting sets as close as possible to the relations in the total set. The final training and test sets include 234 and 58 compounds, respectively.

7.7.Predictivity - Statistics obtained by external validation:

Test set: n = 58, Accuracy 84%; FP rate 41%; FN rate 5%; PPV 85%; NPV 83%; Sensitivity 95%; Specificity 59%; Nb unpredicted 9.

Test set in AD: n= 28, Accuracy 96%, TP 23, TN 4, FP 1, FN 0, PPV 96%, NPV 100%, Sensitivity 100%, Specificity 80%

Test set “could be out of AD”: n = 12, TP 6, TN 1, FP 4, FN 1, PPV 60%, NPV 50%, Sensitivity 86%, Specificity 20%

Test set out of AD: n = 18, TP 10, TN 5, FN 1, FP 5, PPV 83%, NPV 83%, Sensitivity 90%, Specificity 71%

7.8.Predictivity - Assessment of the external validation set:

NA

7.9.Comments on the external validation of the model:

NA

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

No assumption on the mechanism is done.

No rules have been so far reported in the literature. Several mechanisms are expected to be involved, due to the complexity of the endpoint.

8.2.A priori or a posteriori mechanistic interpretation:

NA

8.3.Other information about the mechanistic interpretation:

NA

9.Miscellaneous information

9.1.Comments:

NA

9.2.Bibliography:

[1] V.C. Arena, N.B. Sussman, S. Mazumdar, S. Yu, O.T. Macina (2004) The utility of structure-activity relationship (SAR) models for prediction and covariate selection in developmental toxicity: comparative analysis of logistic regression and decision tree models, SAR QSAR Env. Res.,15, 1-18

[2] Briggs GG, Freeman RK, Yaffe SJ: Drugs in Pregnancy and Lactation Baltimore, MD: Williams and Wilkens, 3 1990

[3] Shepard TH: Catalog of Teratologic Agents Baltimore, MD: John Hopkins University Press, 5 1992

[4] Schüürmann G, Kühne R, Kleint F, Ebert RU, Rothenbacher C, Herth P: A software system for automatic chemical property estimation from molecular structure. Quantitative Structure-Activity Relationships in Environmental Sciences VII. Pensacola, FL: SETAC PressChen F, Schüürmann G 1997, 93-114.

[5] Schüürmann G, Ebert RU, Nendza M, Dearden LC, Paschke A, Kühne R: Prediction of fate-related compound properties. Risk Assessment of Chemicals. An Introduction Dordrecht, NL: Springer Sciencevan Leeuwen K, Vermeire T 2007, 375-426

[6] Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001). https://doi.org/10.1023/A:1010933404324

[7] Floris, M., Manganaro, A., Nicolotti, O. et al. A generalizable definition of chemical similarity for read-across. J Cheminform 6, 39 (2014). https://doi.org/10.1186/s13321-014-0039-1 http:///www.journal.chemistrycentral.com/content/s/S1/S4

9.3.Supporting information:

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

All available dataset are present in the model inside the VEGA software.

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

To be entered by JRC

10.2.Publication date:

To be entered by JRC

10.3.Keywords:

To be entered by JRC

10.4.Comments:

To be entered by JRC