	QMRF identifier (JRC Inventory): **To be entered by JRC**
	QMRF Title: **Developmental/Reproductive Toxicity library (PG) (version 1.1.2)**
	Printing Date: Apr 1**, 2022**

1.QSAR identifier

1.1. QSAR identifier (title):

Developmental/Reproductive Toxicity library (PG) (version 1.1.2)

1.2. Other related models:

The model is a re-implementation of Wu S at al model.

1.2. 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:

25/03/2022

2.2. QMRF author(s) and contact details:

Emilio Benfenati IRCCS-Istituto di Ricerche Farmacologiche Mario Negri Via La Masa 19, 20156Milano, Italy emilio.benfenati@marionegri.it https://www.marionegri.it/

2.3. Date of QMRF update(s):

NA

2.4. QMRF update(s):

NA

2.5. Model developer(s) and contact details:

[1] Alberto Manganaro RCCS-Istituto di Ricerche Farmacologiche Mario Negri Via La Masa 19,20156 Milano, Italy alberto.manganaro@marionegri.it

[2] Marco Marzo Istituto di ricerche farmacologiche Mario Negri - IRCSS marco.marzo@marionegri.it

2.6. Date of model development and/or publication:

28/09/2016

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

[1] Wu S, Fisher J, Naciff J, Laufersweiler M, Lester C, Daston G, Blackburn K. Framework for identifying chemicals with structural features associated with the potential to act as developmental or reproductive toxicants. Chem Res Toxicol. 2013 Dec 16;26(12):1840-61. doi: 10.1021/tx400226u

Other references related to:

[2] Marzo M, Kulkarni S, Manganaro A, Roncaglioni A, Wu S, Barton-Maclaren TS, Lester C, Benfenati E. Integrating in silico models to enhance predictivity for developmental toxicity. Toxicology. 2016 Aug 31;370:127-137. doi: 10.1016/j.tox.2016.09.015

[3] Benfenati E, Manganaro A, Gini G

VEGA-QSAR: AI inside a platform for predictive toxicology

Proceedings of the workshop "Popularize Artificial Intelligence 2013", December 5th 2013, Turin, Italy

Published on CEUR Workshop Proceedings Vol-1107

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:

Homo sapiens (male and female)

3.2. Endpoint:

Developmental and Reproductive toxicity.

3.3. Comment on endpoint:

Developmental and reproductive toxicity model classification:
a) **Receptor interaction of DART- relevance** (5 receptor groups/10 receptors mentioned)
&
b) **20 Structural Alerts related to DART**
(incl. 1 group: inorganics & metallic comp. / 19 org. groups
(all specified in (2)): related to *in vivo* incl. human evidence – the nature of this evidence of concern for DART is only specified in refs. incl. in the ref. list of (2))

3.4. Endpoint units:

Model is a classification so there are no units (Adimensional)

3.5. Dependent variable:

Binary classification (DEV toxicant, DEV NON-toxicant, note that DEV non toxicant includes evidence for lack of DEV toxicity as well as lack of evidence)

3.6. Experimental protocol:

Data are extracted from literature [1]

3.7. Endpoint data quality and variability:

Data collection is described in "A Framework for Identifying Chemicals with Structural Features Associated with Potential to Act as Developmental or Reproductive Toxicants" Wu et al. 2013. (DOI:10.1021/tx400226u). The final dataset counts 685 substances: from the original dataset (n. 716) we selected substances on the basis of their structure (e.g. polymers, inorganics compounds and organometals were excluded) and with data for at least one endpoint (developmental toxicity, reproductive toxicity).

For this model, experimental values of data are labeled as:

Developmental NON-toxicant and no data on reproductive toxicity

Developmental Toxicant and no data available on reproductive toxicity

Developmental Toxicant but reproductive NON-toxicant

Both reproductive and developmental NON-toxicant

Both reproductive and developmental Toxicant

Reproductive NON-Toxicant and no data on developmental toxicity

Reproductive toxicant and no data available on developmental toxicity

Reproductive toxicant but developmental NON-toxicant

4.Defining the algorithm - OECD Principle 2

4.1. Type of model:

The P&G model is a decisional tree with six nodes representing different chemical features. there are 25 categories as further splitting of these six nodes. Each category represents

groups of compounds with a determined biological activity or common chemical feature. The 25 categories are further divided into 129 chemical subcategories or “rules” defined by a structural backbone or scaffold

with substituents. Considering all possible substituent positions, enumeration of the 129 subcategories results in a library of ~ 185,000 structures (mono-constituent organic substances).

4.2. Explicit algorithm:

The model evaluates the applicability domain against a number of similar chemicals, in an automatic read across approach, using the applicability domain index of VEGA (c.f. section 5.1 and 5.2 below). It is based on a similarity check to compare the queried substances with those used to develop the model and to verify how accurate their predicted values are. These similar compounds are selected amongst mono-constituent organic substances included in the categories and subcategories of the decisional tree.

4.3. Descriptors in the model:

The model is a structure-based model and does not make use of descriptors

4.4. Descriptor selection:

NA

4.5. Algorithm and descriptor generation:

NA

4.6. Software name and version for descriptor generation:

NA

4.7. Chemicals/Descriptors ratio:

NA

5.Defining the applicability domain - OECD Principle 3

5.1. Description of the applicability domain of the model:

The 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 (c.f. below section 5.2 where it is described how the ADI is calculated for this model)

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

If 1 ≥ AD index > 0.9, the predicted substance is regarded in the Applicability Domain of the model. It corresponds to “good reliability” of prediction.

If 0.9 ≥ AD index > 0.65, the predicted substance could be out of the Applicability Domain of the model. It corresponds to “moderate reliability” of prediction.

If AD index ≤ 0.65, the predicted substance is regarded out of the Applicability Domain of the model. It corresponds to “low reliability” of prediction.

Indices are calculated on the first k = 2 most similar molecules, each having S_k similarity value with the target molecule. For the purpose of classification statistics, all different active (toxicant) classes are mapped into a unique active class, versus the single non-toxicant class.

Similarity index (*IdxSimilarity*) is calculated as:

$$\frac{\sum_k S_k}{k} \times (1 - Diam^2)$$

where *Diam* is the difference in similarity values between the most similar molecule and the *k*-th molecule.

Accuracy index (*IdxAccuracy*) is calculated as:

$$\frac{\sum_c log\,(1 + S_c)}{\sum_k log\,(1 + S_k)}$$

where the molecules with *c* index are the subset of the *k* molecules where the prediction of the model matches with the experimental value of the molecule.

Concordance index (*IdxConcordance*) is calculated as:

$$\frac{\sum_c log\,(1+S_c)}{\sum_k log\,(1+S_k)}$$

where the molecules with *c* index are the subset of the *k* molecules where the experimental value of the molecule matches with the prediction made for the target molecule.

ACF contribution (*IdxACF*) index is calculated as

$$ACF = rare \times missing$$

where: *rare* is calculated on the number of fragments found in the molecule and found in the training set in less than 3 occurences as following: if the number is 0, *rare* is set to 1.0; if the number is 1, *rare* is set to 0.6; otherwise *rare* is set to 0.4

missing is calculated on the number of fragments found in the molecule and never found in the training set as following: if the number is 0, *missing* is set to 1.0; if the number is 1, *missing* is set to 0.6; otherwise *missing* is set to 0.4

AD final index is calculated as following:

$$ADI = (IdxSimilarity^{0.5} \times IdxAccuracy^{0.25} \times IdxConcordance^{0.25}) \times IdxACF$$

5.2. Method used to assess the applicability domain:

The chemical similarity is measured with the algorithm developed for VEGA. Full details in the VEGA website (www.vegahub.eu), including the open access paper describing it [3]. 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 > 0.85, strongly similar compounds with known experimental value in the training set have been found

If 0.85 ≥ index > 0.75, only moderately similar compounds with known experimental value in the training set have been found

If index ≤ 0.75, 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 index < 0.5, accuracy of prediction for similar molecules found in the training set is good

If 0.85 > index ≥0.5, accuracy of prediction for similar molecules found in the training set is not optimal

If index ≥ 0.85, 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 index < 0.5, molecules found in the training set have experimental values that agree with the target compound predicted value

If 0.85 > index ≥ 0.5, similar molecules found in the training set have experimental values that slightly disagree with the target compound predicted value

If index ≥ 0.85, 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

5.3. Software name and version for applicability domain assessment:

VEGA (www.vegahub.eu)

5.4. Limits of applicability:

VEGA provides a quantitative value for the prediction of each substance. This helps the user to identify potential critical aspects, which are indicated. Similar compounds are shown.

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: No

Smiles: Yes

Formula: No

INChI: No

MOL file: No

NanoMaterial: No

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

No

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

No

6.5. Other information about the training set:

The dataset was composed of 685 substances retrieved from "A Framework for Identifying Chemicals with Structural Features Associated with Potential to Act as Developmental or Reproductive Toxicants" Wu et al. 2013. (DOI:10.1021/tx400226u). Data set in also present as supporting material of the paper.

6.6. Pre-processing of data before modelling:

NA

6.7. Statistics for goodness-of-fit:

Sensitivity 89%, Specificity 44%, Accuracy 85%, MCC 0.27

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

NA

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

NA

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:

NO c.f. however reference ((2)

7.2. Available information for the external validation set:

c.f. below point 7.7

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

NA

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

NA

7.5. Other information about the external validation set:

NA

7.6. Experimental design of test set:

NA

7.7. Predictivity - Statistics obtained by external validation:

In ref. 2 the following is mentioned: Test sets
(**NB: some overlaps with the training set occurred but is not specified in ref. (2)**):
LoDS (2008) & vet. medicine: 106+, sensitivity: 74%
CAESAR-list: 194+, sensitivity 89 %
RIVM-list: 110+, sensitivity: 88%
LoDS: EU Harmonized List of Classified Dangerous Substances (EU 2008)
The three above mentioned lists contain exclusively substances regarded as developmental/ reproductive toxicants (c.f. ref. (2))

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:

NA

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:

Due to the low specificity and the nature of the model only positive predictions should be taken as indications only (c.f. section e.g. like positive evidence from OECD QSAR TB profilers. Note therefore that a positive prediction is *not a QSAR prediction: use only pos. calls as indicative"* or "*Use only pos. results as indications together with other relevant information pointing in the same direction"*
→ Useful if pos as an indication of potential concern for DART

9.2. Bibliography:

[1] Marzo M, Kulkarni S, Manganaro A, Roncaglioni A, Wu S, Barton-Maclaren TS, Lester C, Benfenati E. Integrating in silico models to enhance predictivity for developmental toxicity. Toxicology. 2016 Aug 31;370:127-137. doi: 10.1016/j.tox.2016.09.015

[2] Wu S, Fisher J, Naciff J, Laufersweiler M, Lester C, Daston G, Blackburn K. Framework for identifying chemicals with structural features associated with the potential to act as developmental or reproductive toxicants. Chem Res Toxicol. 2013 Dec 16;26(12):1840-61. doi: 10.1021/tx400226u

[3] Floris et al. "A generalizable definition of chemical similarity for read-across." Journal of cheminformatics 6.1 (2014): 39

9.3. Supporting information:

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

All available datasets 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