	QMRF identifier (JRC Inventory): **To be entered by JRC**
	QMRF Title: **Thyroid Receptor Beta effect (NRMEA) - v. 1.0.0**
	Printing Date: **14-feb-2020**

1.QSAR identifier

1.1.QSAR identifier (title):

Thyroid Receptor Beta effect (NRMEA) - v. 1.0.0

1.2.Other related models:

This is the description of the VEGA model that implements the "Thyroid Receptor Beta effect (NRMEA) - v. 1.0.0" developed by the group of Prof. Wei Shi, University of Nanjing

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:

14-02-2020

2.2.QMRF author(s) and contact details:

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

Marco Marzo Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy marco.marzo@marionegri.it https://www.marionegri.it/

2.3.Date of QMRF update(s):

No update

2.4.QMRF update(s):

No update

2.5.Model developer(s) and contact details:

[1]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]Tan Haoyue State Key Laboratory of Pollution Control and Resources Reuse, School of the Environment, Nanjing University, Nanjing, Jiangsu 210023, China Tan Haoyue Tan Haoyue 417695798@qq.com 417695798@qq.com

[3]Wei Shi State Key Laboratory of Pollution Control and Resources Reuse, School of the Environment, Nanjing University, Nanjing, Jiangsu 210023, China Tan Haoyue Wei Shi njushiwei@nju.edu.cn 417695798@qq.com

2.6.Date of model development and/or publication:

24 sept 2019

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 the training, test and validation sets are available.

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

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Homo sapiens

3.2.Endpoint:

Endocrine disrupting chemicals Nuclear receptor-mediated endocrine disruption Thyroid Receptor Beta effect

3.3.Comment on endpoint:

The effect of compound on Thyroid receptor beta (classification)

3.4.Endpoint units:

Model is a classification so there is no units, possible results should be Active/Inactive prediction and Agonist/A-Anta/Antagonist.

3.5.Dependent variable:

The dependent variable is Endocrine effect, as binary classification: 0 (inactive), 1 (active).

3.6.Experimental protocol:

Model are built based on two in vitro assays: i) reporter gene assay and ii) cytotoxicity.

3.7.Endpoint data quality and variability:

Chemical data were collected from two open free libraries, including ToxCast/Tox21 (https://www.epa.gov/chemical-research/toxcast-dashboard) and ChEMBL (https://www.ebi.ac.uk/chembl/).

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

Hierarchy featured fragments-based model.

4.2.Explicit algorithm:

Hierarchy fragment features

model is a hierarchical tree on three levels: fragments on first and second levels define if a compound is active on receptor, third fragments define the type of activity of the compounds on the receptor

4.3.Descriptors in the model:

Structural fragment based. The model is a structure-based model and does not make use of Descriptors.

4.4.Descriptor selection:

Structure alerts selected are described in the guide associated at the model present in the Vega platform.

4.5.Algorithm and descriptor generation:

Firstly, we took advantage of primary fragments to select compounds have the possibility of being active. Secondly, compounds were screened by secondary fragments according to their different types of primary fragments. Each compound contains primary and secondary fragments simultaneously, would be predicted as active compounds. Next, the types of disrupting activities of identified active compounds would be distinguished based on their tertiary fragments. This hierarchy fragments allow the exploration of a wide range of structurally diverse active compounds which could bind to receptors, and build up an array of rules focusing on Agonist/A-Anta/Antagonist disruption.

Structural fragments are calculating using substructure frequency analysis and substructure percentage analysis, and PubChem fingerprint database was used to generate primary fragments to describe active disruptors as much as possible. According to the structural features, chemicals were divided into several subtypes based on their primary fragments. Then, specific secondary fragments were generated by using SARpy software (https://www.vegahub.eu/portfolio-item/sarpy/). Active disruptors, containing primary and related secondary fragments, were used to extracted disrupting tertiary fragments. Disruptors were clustered according to their primary and secondary fragments, then, tertiary fragments were extracted independently by using SARpy software.

4.6.Software name and version for descriptor generation:

Padel-descriptor

Padel-descriptor (for primary fragment) https://www.softpedia.com/get/Science-CAD/PaDEL-Descriptor.shtml

SARpy software

SARpy software (for secondary and tertiary fragment) https://www.vegahub.eu/portfolio-item/sarpy/

4.7.Chemicals/Descriptors ratio:

5462 chemicals/ 37 structural alerts

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.

5.2.Method used to assess the applicability domain:

The chemical similarity is measured with the algorithm developed for VEGA. Full details are in the VEGA website (www.vegahub.eu), including the open access paper describing it. 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.

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: No

Smiles: Yes

Formula: No

INChI: No

MOL file: No

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:

6.6.Pre-processing of data before modelling:

All of the chemical data were prepared by removing all false SMILES strings, and deleting all duplicate compounds. Then, the data set was split randomly into training set and test set in the ratio of 4:1 (KNIME Analytics Platform, https://www.knime.com/). The training set was used to provide feature fragments and local physicochemical properties present across the completely active set. The test set was used in the external validation.

6.7.Statistics for goodness-of-fit:

We Active/Inactive prediction

Training set: accuracy: 0.995; sensitivity: 0.83; specificity: 0.998

Agonist/A-Anta/Antagonist prediction

Training set accuracy:0.98

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

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:

No

7.2.Available information for the external validation set:

CAS RN: No

Chemical Name: No

Smiles: No

Formula: No

INChI: No

MOL file: No

NanoMaterial: No

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:

7.6.Experimental design of test set:

7.7.Predictivity - Statistics obtained by external validation:

Active/Inactive prediction

Test set: accuracy: 0.99; sensitivity: 0.71; specificity: 0.997

Agonist/A-Anta/Antagonist prediction

test set accuracy: 1

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:

With NRMEA, you can predict chemicals Thyroid Receptor Beta effect.

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]Bruning JB, Parent AA, Gil G, Zhao M, Nowak J, Pace MC, et al. 2010. Coupling of receptor conformation and ligand orientation determine graded activity. Nature Chemical Biology 6:837-843.

[2]Brust R, Shang J, Fuhrmann J, Mosure SA, Bass J, Cano A, et al. 2018. A structural mechanism for directing corepressor-selective inverse agonism of ppar. Nature Communications 9:4687.

[3]Chen Q, Wang X, Wei S, Yu H, Zhang X, Giesy JP. 2016. Identification of thyroid hormone disruptors among ho-pbdes: In vitro investigations and co-regulator involved simulations. Environmental Science & Technology 50.

[4]Chen Q, Tan H, Yu H, Wei S. 2018. Activation of steroid hormone receptors: Shed light on the in silico evaluation of endocrine disrupting chemicals. Science of the Total Environment 631-632:27-39.

[5]Delfosse V, Grimaldi M, Pons J-L, Boulahtouf A, le Maire A, Cavailles V, et al. 2012. Structural and mechanistic insights into bisphenols action provide guidelines for risk assessment and discovery of bisphenol a substitutes. Proceedings of the National Academy of Sciences 109:14930-14935.

[6]Eva EPÁ, Arnold LA, Arnold AA, Phuong N, Edson Delgado R, Ellena M, et al. 2007. A surface on the androgen receptor that allosterically regulates coactivator binding. Proceedings of the National Academy of Sciences of the United States of America 104:16074-16079.

[7]Kleinstreuer NC, Ceger P, Watt ED, Martin M, Houck K, Browne P, et al. 2016. Development and validation of a computational model for androgen receptor activity. Chemical Research in Toxicology 30:946.

[8]Li Y, Perera L, Coons LA, Burns KA, Tyler RJ, Pelch KE, et al. 2018. Differential in vitro biological action, coregulator interactions, and molecular dynamic analysis of bisphenol a (bpa), bpaf, and bps ligand-er complexes. Environmental Health Perspectives 126:017012.

[9]Mellor CL, Steinmetz FP, Cronin MT. 2016. Using molecular initiating events to develop a structural alert based screening workflow for nuclear receptor ligands associated with hepatic steatosis. Chemical Research in Toxicology 29:203.

[10]Patience B, Judson RS, Casey WM, Kleinstreuer NC, Thomas RS. 2015. Screening chemicals for estrogen receptor bioactivity using a computational model. Environmental Science & Technology 49:8804-8814.

[11]Steinmetz FP, Mellor CL, Meinl T, Cronin MT. 2015. Screening chemicals for receptor-mediated toxicological and pharmacological endpoints: Using public data to build screening tools within a knime workflow. Qsar & Combinatorial Science 34:171-178.
[12]Vanessa D, Marina G, Vincent C, Patrick B, William B. 2014. Structural and functional profiling of environmental ligands for estrogen receptors. Environ Health Perspect 122:1306-1313.
[13]Wang X, Zhang X, Xia P, Zhang J, Wang Y, Zhang R, et al. 2017. A high-throughput, computational system to predict if environmental contaminants can bind to human nuclear receptors. Science of the Total Environment 576:609-616.
[14]Wu Y, Doering JA, Ma Z, Tang S, Liu H, Zhang X, et al. 2016. Identification of androgen receptor antagonists: Invitro investigation and classification methodology for flavonoid. Chemosphere 158:72-79.

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