

	QMRF identifier (JRC Inventory): To be entered by JRC
	QMRF Title: Nuclear Receptor-mediated Endocrine Activity (NRMEA)
	Printing Date: 2019-5-14

1. QSAR identifier

1.1. QSAR identifier (title):

Nuclear Receptor-mediated Endocrine Activity (NRMEA)

1.2. Other related models:

1.3. Software coding the model:

NRMEA v 1.1.1

The model provides a qualitative prediction of 12 classical nuclear receptor-mediated endocrine disruptions, including Androgen Receptor (AR), Estrogen Receptor a/b (ERa/b), Glucocorticoid Receptor (GR), Mineralocorticoid Receptor (MR), Progesterone Receptor (PR), Retinoic Acid Receptor a/b/g (RARa/b/g), Thyroid Hormone Receptor a/b (TRa/b), and Vitamin D Receptor (VDR). It is implemented inside the VEGA online platform, accessible at: <https://www.vegahub.eu>.

Nuclear Receptor-mediated Endocrine Activity

2. General information

2.1. Date of QMRF:

13 May 2019

2.2. QMRF author(s) and contact details:

Tan Haoyue, Shi Wei, Chen Qinchang, Yu Hongxia State Key Laboratory of Pollution Control and Resources Reuse, School of the Environment, Nanjing University, Nanjing, Jiangsu 210023, China
Tan Haoyue: 417695798@qq.com; Wei Shi: njushiwei@nju.edu.cn

2.3. Date of QMRF update(s):

13 May 2019

2.4. QMRF update(s):

13 May 2019

2.5. Model developer(s) and contact details:

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

2.6. Date of model development and/or publication:

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

2.8. Availability of information about the model:

Prediction interface and validation results available at
<https://www.vegahub.eu>.

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

3. Defining the endpoint - OECD Principle 1

3.1. Species:

Homo sapiens (Active/Inactive prediction and
Agonist/A-Anta/Antagonist disrupting prediction) about potential
disruptors for 12 classical nuclear receptors by competitive binding and
receptor gene steps of MIE.

3.2.Endpoint:

Endocrine disrupting chemicals Nuclear receptor-mediated endocrine disruption

3.3.Comment on endpoint:

It's worth noting that three predictive models (AR, ERa, GR) are built based on three in vitro assays: i) competitive binding assay, ii) reporter gene assay and iii) cytotoxicity. Whereas other predictive models (ERb, PR, MR, RARa/b/g, TRa/b, VDR) are built based on two in vitro assays: i) reporter gene assay and ii) cytotoxicity. Therefore, the predicted results of first three models are under the classical nuclear genomic mechanism, in which ligand-NR binding directly and leading abnormal transcription, and the later couldn't predict the disrupting mechanism result from ligand-bound NRs.

3.4.Endpoint units:

Active/Inactive prediction and Agonist/A-Anta/Antagonist
disrupting prediction (classification)

3.5.Dependent variable:

Endocrine disruption

3.6.Experimental protocol:

3.7.Endpoint data quality and variability:

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

Hierarchy featured fragments-based prediction model

4.2.Explicit algorithm:

Definition and description available at <https://www.vegahub.eu>.

4.3.Descriptors in the model:

Hierarchy featured fragments, including primary fragment, secondary fragment and tertiary fragment.

4.4.Descriptor selection:

Primary fragment is extracted by Pubchem fingerprinter by using substructure frequency analysis and substructure percentage analysis, secondary fragment and tertiary fragment are extracted by using SARpy software, all of which are available at <https://www.vegahub.eu>.

4.5.Algorithm and descriptor generation:

4.6.Software name and version for descriptor generation:

Padel-descriptor (for primary fragment), SARpy software (for secondary and tertiary fragment)
<https://www.softpedia.com/get/Science-CAD/PaDEL-Descriptor.shtml>;
<https://www.vegahub.eu/portfolio-item/sarpy/>

4.7.Chemicals/Descriptors ratio:

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

NRMEA is built on the base of hierarchy featured fragments, thus diverse compounds can be predicted except for mixtures. The details are available at <https://www.vegahub.eu>.

5.2.Method used to assess the applicability domain:

5.3. Software name and version for applicability domain assessment:

5.4. Limits of applicability:

6. Internal validation - OECD Principle 4

6.1. Availability of the training set:

Yes

6.2. Available information for the training set:

yes

CAS RN: Yes

Chemical Name: Yes

Smiles: Yes

Formula: Yes

INChI: Yes

MOL file: Yes

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:

Datasets available at <https://www.vegahub.eu>.

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:

The performance of NRMEA are available at <https://www.vegahub.eu>.

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:

Yes

7.2. Available information for the external validation set:

yes

CAS RN: Yes

Chemical Name: Yes

Smiles: Yes

Formula: Yes

INChI: Yes

MOL file: Yes

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:

7.6.Experimental design of test set:

7.7.Predictivity - Statistics obtained by external validation:

The performance of NRMEA are available at <https://www.vegahub.eu>.

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 without disrupting data about 12 NRs-mediated disrupting activities aforementioned. 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/>).

It's worth noting that three predictive models (AR, ERa, GR) are built based on three in vitro assays: i) competitive binding assay, ii) reporter gene assay and iii) cytotoxicity. Whereas other predictive models (ERb, PR, MR, RARa/b/g, TRa/b, VDR) are built based on two in vitro assays: i) reporter gene assay and ii) cytotoxicity. Therefore, the predicted results of first three models are under the classical nuclear genomic mechanism, in which ligand-NR binding directly and leading abnormal transcription, and the later couldn't predict the disrupting mechanism result from ligand-bound NRs .

8.2.A priori or a posteriori mechanistic interpretation:

Compounds with the same primary and secondary fragments, will induce the same ligand-receptor interactions. Then, tertiary will affect the coregulator recruitment of ligand-receptor complex, which is critical for transcriptional activity.

8.3.Other information about the mechanistic interpretation:

Criterion 1: compounds have no primary fragment are inactive

Criterion 2: active compounds have at least one secondary fragment and inactive compounds have none

Criterion 3: tertiary fragments decide the disrupting activities of active compounds

9.Miscellaneous information

9.1.Comments:

Public model interface: <https://www.vegahub.eu>.

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, Meini 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)

AR database.xlsx	file:///H:/H\\Prediction Model\12NRs database\AR database.xlsx
ERa database.xlsx	file:///H:/H\\Prediction Model\12NRs database\ERa database.xlsx
ERb database.xlsx	file:///H:/H\\Prediction Model\12NRs database\ERb database.xlsx

GR database.xlsx	file:///H:\H\Prediction Model\12NRs database\GR database.xlsx
MR database.xlsx	file:///H:\H\Prediction Model\12NRs database\MR database.xlsx
PR database.xlsx	file:///H:\H\Prediction Model\12NRs database\PR database.xlsx
RARa database.xlsx	file:///H:\H\Prediction Model\12NRs database\RARa database.xlsx
RARb database.xlsx	file:///H:\H\Prediction Model\12NRs database\RARb database.xlsx
RARg database.xlsx	file:///H:\H\Prediction Model\12NRs database\RARg database.xlsx
TRa database.xlsx	file:///H:\H\Prediction Model\12NRs database\TRa database.xlsx
TRb database.xlsx	file:///H:\H\Prediction Model\12NRs database\TRb database.xlsx
VDR database.xlsx	file:///H:\H\Prediction Model\12NRs database\VDR database.xlsx

Test set(s) Supporting information

Information about training set and test set.docx	file:///C:\Users\tanhaoyue\Desktop\Information about training set and test set.docx
--	---

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