

	QMRF identifier (JRC Inventory): To be entered by JRC
	QMRF Title: Mitochondrial Dysfunction Model V1.0
	Printing Date: April 3 2024

1. QSAR identifier

1.1. QSAR identifier (title):

Mitochondrial Dysfunction Model V1.0

1.2. Other related models:

- Inhibition Mitochondrial Complexes Model V1.0
- Oxidative stress Model V1.0

1.3. Software coding the model:

The alternative cloud Platform <https://platform.alternative-project.eu/>

2. General information

2.0. Abstract:

2.1. Date of QMRF:

April 2024

2.2. QMRF author(s) and contact details:

Edoardo Luca Viganò – Istituto di ricerche farmacologiche Mario Negri – edoardo.vigano@marionegri.it

2.3. Date of QMRF update(s):

NA

2.4. QMRF update(s):

NA

2.5. Model developer(s) and contact details:

Edoardo Luca Viganò – Istituto di ricerche farmacologiche Mario Negri – edoardo.vigano@marionegri.it

2.6. Date of model development and/or publication:

2024

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

[1] <https://doi.org/10.3390/toxics12010087>

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:

Cell line assays.

Assays selected to evaluate the mitochondrial dysfunction: APR HepG2 MitoMass 24 h dn, APR HepG2 MitoMass 24 h up, APR HepG2 MitoMass 72 h dn, APR HepG2 MitoMass 72 h up, APR HepG2 MitoMembPot 24 h dn, APR HepG2 MitoMembPot 24 h up, APR HepG2 MitoMembPot 72 h dn, APR HepG2 MitoMembPot 72 h up, ATG XTT Cytotoxicity up, TOX21 MMP ratio down, TOX21 MMP ratio up, TOX21 MMP rhodamine.

3.2.Endpoint:

Mitochondrial dysfunction

3.3.Comment on endpoint:

Mitochondrial dysfunction has emerged as a critical area of investigation due to its implications for numerous diseases and physiological processes. Indeed, mitochondria play a central role in cell metabolism, serving as fundamental components in energy production, metabolism, and cellular signaling. Dysfunction in these organelles can lead to a wide range of health issues, including neurodegenerative diseases, and metabolic, liver, and cardiac disorders.

The data are labeled as active or inactive for classification modeling purposes. We define activity as follows: a chemical is considered active if it shows a hit call label as active in at least one of the selected assays; otherwise, it is labeled as inactive. The individual assay label was available in the files downloaded from the ICE platform. In particular, raw data provided by a vendor or laboratory underwent processing, indexing, transformation, and normalization using standardized methods. Subsequently, the concentration–response data are subjected to modeling through three selected models (constant, Hill, and gain–loss). If any models fit sufficiently, the chemical–assay pair is considered ‘active’ (hit call = active); otherwise, the final hit call is ‘inactive’.

3.4. Endpoint units:

Adimensional

3.5. Dependent variable:

The dependent variable is mitochondrial dysfunction, as binary classification: 0 (non-inhibitor), 1 (inhibitor)

3.6. Experimental protocol:

We considered the results of different biological assays reported in ICE database (<https://ice.ntp.niehs.nih.gov/>, accessed on 15 October 2023), which provides high-quality curated data to support the development and evaluation of new, revised, and alternative methods.

3.7. Endpoint data quality and variability:

The data was retrieved ICE database [2], which provides high-quality curated data to support the development and evaluation of new, revised, and alternative methods. The initial dataset has curated and cleaned of all ambiguous or mixed structures, polymers, inorganic compounds, metallo-organic compounds, salts, complexes and we maintain all the QSAR ready structures. The final data set consist of 232 chemicals.

4. Defining the algorithm - OECD Principle 2

4.1. Type of model:

Neural Languages Process models (NLP).

Deep Neural Network that uses as input the SMILES embedded.

4.2. Explicit algorithm:

The architecture of these networks can process chemical notation directly as text using specific layers for text vectorization and character embedding, which describes each character or segment of the string as a numerical vector. The model is designed to handle sequential data and capture dependencies between tokens in the input sequence and the properties.

4.3. Descriptors in the model:

Molecular descriptors are not used. Instead, the models employ the semantic and grammatical concepts behind SMILES notations to encode chemical information in a suitable manner for modeling.

4.4. Descriptor selection:

na

4.5. Algorithm and descriptor generation:

The Neural networks model is based on learning encoding. That means the model generates hidden representations that capture the semantic information of the input that in our case is the SMILES notation.

4.6. Software name and version for descriptor generation:

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 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 $1 \geq \text{AD index} > 0.80$, the predicted substance is regarded in the Applicability Domain of the model. It corresponds to "good reliability" of prediction.

If $0.80 \geq \text{AD index} > 0.6$, the predicted substance could be out of the Applicability Domain of the model. It corresponds to "moderate reliability" of prediction.

If $\text{AD index} \leq 0.6$, 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 Applicability domain 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 [1]. The VEGA 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 $1 \geq \text{index} > 0.80$, strongly similar compounds with known experimental value in the training set have been found.

If $0.80 \geq \text{index} > 0.6$, only moderately similar compounds with known experimental value in the training set have been found.

If $\text{index} \leq 0.6$, 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 $\text{index} < 0.6$, accuracy of prediction for similar molecules found in the training set is good

If $0.8 > \text{index} \geq 0.6$, accuracy of prediction for similar molecules found in the training set is not optimal

If $\text{index} \geq 0.8$, 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 $\text{index} < 0.6$, molecules found in the training set have experimental values that agree with the target compound predicted value

If $0.8 > \text{index} \geq 0.6$, similar molecules found in the training set have experimental values that slightly disagree with the target compound predicted value

If $\text{index} \geq 0.8$, 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: a first index RARE takes into account rare fragments (those who occur less than three times in the training set), having value of 1 if no such fragments are found, 0.85 if up to 2 fragments are found, 0.7 if more than 2 fragments are found; a second index NOTFOUND takes into account not found fragments, having value of 1 if no such fragments are found, 0.6 if a fragments is found, 0.4 if more than 1 fragment is found. Then, the final index is given as the product $\text{RARE} * \text{NOTFOUND}$. Defined intervals are:

If $\text{index} = 1$, all atom centered fragment of the compound have been found in the compounds of the training set

If $1 > \text{index} \geq 0.7$, some atom centered fragment of the compound have not been found in the compounds of the training set or are rare fragments

If $\text{index} < 0.7$, 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:

Rdkit Python packages

5.4. Limits of applicability:

The model is not applicable to inorganic chemicals and salts.

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

Smiles: Yes

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

no

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

All

6.5. Other information about the training set:

n 5004

6.6. Pre-processing of data before modelling:

The SMILES notation is retrieved from the most reliable databases starting from CASRN.

The SMILES retrieved are then curated using the preprocess methods developed by winter at all [4].

6.7. Statistics for goodness-of-fit:

Training set: n = 4504,

Balanced Accuracy: 0.886

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:

Yes

7.2. Available information for the external validation set:

CAS RN: Yes

Smiles: 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:

n: 500

7.6. Experimental design of test set:

NA

7.7. Predictivity - Statistics obtained by external validation:

Test set: n = 500

Balanced Accuracy 0.82

Precision 0.62

Sensitivity 0.78

Specificity 0.86

MCC 0.59

F1-Score: 0.69

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:

NA

9.2. Bibliography:

[1] Viganò EL, Ballabio D, Roncaglioni A. Artificial Intelligence and Machine Learning Methods to Evaluate Cardiotoxicity following the Adverse Outcome Pathway Frameworks. *Toxics*. 2024; 12(1):87.

<https://doi.org/10.3390/toxics12010087>

[2] ICE database (<https://ice.ntp.niehs.nih.gov/>, accessed on 15 October 2023)

[3] Krishna, S.; Berridge, B.; Kleinstreuer, N. High-Throughput Screening to Identify Chemical Cardiotoxic Potential. *Chem. Res. Toxicol.* 2021, 34, 566–583.

[4] Winter, R.; Montanari, F.; Noé, F.; Clevert, D.-A. Learning Continuous and Data-Driven Molecular Descriptors by Translating Equivalent Chemical Representations. *Chem. Sci.* 2019, 10, 1692–1701.

9.3. Supporting information:

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

All available datasets are present in alternative cloud platform and will be in the model inside the VEGA software.