	***QMRF identifier (JRC Inventory):*** **To be entered by JRC**
	QMRF Title: **Skin Sensitization model (IRFMN/JRC) 1.0.1**
	Printing Date: Nov 6, 2022

1.QSAR identifier

1.1.QSAR identifier (title):

Skin Sensitization model (IRFMN/JRC) 1.0.1

1.2.Other related models:

Skin sensitization model (CAESAR) version 2.1.7

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:

Nov 2022

2.2.QMRF author(s) and contact details:

Gianluca Selvestrel Istituto di Ricerche farmacologiche Mario negri – IRCCS gianluca.selvestrel@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 Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2,20156 Milano, Italy alberto.manganaro@marionegri.it https://www.marionegri.it/

[2] Serena Manganelli

[3] Davide Asturiol

2.6.Date of model development and/or publication:

November 2017

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

[1] Selvestrel, G., Robino, F., Baderna, D., Manganelli, S., Asturiol, D., Manganaro, A., Russo, M. Z., Lavado, G., Toma, C., Roncaglioni, A., & Benfenati, E. (2021). Sphera Cosmolife: A new tool for the risk assessment of cosmetic products. ALTEX - Alternatives to Animal Experimentation, 38(4), 565–579. https://doi.org/10.14573/altex.2010221

[2] 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:

Other QMRF for this model is not available

3.Defining the endpoint - OECD Principle 1

3.1.Species:

CBA mice

3.2.Endpoint:

OECD. Test Guideline No. 429: Skin Sensitisation. Local Lymph Node Assay [3]

3.3.Comment on endpoint:

Skin sensitizers are substances able to elicit an allergic response following contact with the skin, termed allergic contact dermatitis (ACD) in humans. The LLNA assay evaluates the induction phase of the allergenic response. Molecules have been classified as "sensitizers" or "non sensitizers"

3.4.Endpoint units:

Adimensional

3.5.Dependent variable:

NA

3.6.Experimental protocol:

Based on the OECD 429 test guideline (2002 (now replaced) & 2010) [3]. The difference between these two TG 429 versions are however judged to be minor.

3.7.Endpoint data quality and variability:

Data on mono-constituent organic substances were collected from the CAESAR model database (209 substances) [1] and from Asturiol et al. (2016) [2] (269 substances). The dataset was split into a training (80%) and a test (20%) set. To partition the chemicals into training set and test set, assuring high diversity and keeping the ratio sensitizer/non-sensitizer, the following procedure was used. The chemicals were initially separated into sensitizer and non-sensitizer. The subsequent steps were carried out separately for sensitizer and non-sensitizer. Each group was clustered based on chemical similarity defined by the chemicals' fingerprints (RDKit atomic pairs) into as many clusters as the number of chemicals divided by 10. Subsequently, 80% of clustered chemicals were assigned randomly to the training set, using the assigned cluster as stratification variable. The remaining 20% of chemicals were assigned to the test set. The chemicals were structurally diverse, and the distribution between sensitizers and non-sensitizers was preserved.

The training set contains 264 compounds. The test set counts 68 compounds

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

Decision trees (rpart R package) based on 2D DRAGON descriptors: nDB, IC1, SIC2, GATS8s, nRCHO, CATS2D_01_NL, F04[C-O], MLOGP

4.2.Explicit algorithm:

Decision tree created using rpart R packages, as encoded in VEGA

4.3.Descriptors in the model:

[1]nDB number of double bonds

[2]IC1 Information Content index (neighborhood symmetry of 1-order)

[3]SIC2 Structural Information Content index (neighborhood symmetry of 2-order)

[4]GATS8s Geary autocorrelation of lag 8 weighted by I-state

[5]nRCHO number of aldehydes (aliphatic)

[6]CATS2D_01_NL CATS2D Negative-Lipophilic at lag 01

[7]F04[C-O] Frequency of C - O at topological distance 4

[8]MLOGP Moriguchi octanol-water partition coeff. (logP)

4.4.Descriptor selection:

An in-house tool developed in the R statistical platform has been used to select the best descriptors set and size to be employed for the final model. The approach was based on a forward selection technique: starting from the descriptor most correlated with the experimental data, at each iteration the descriptor leading to the best model (among all the available descriptors) was added, until the size of 25 descriptors. Models have been built, with a linear discriminant analysis modelling, applied with a bootstrap cross-validation approach (n = 100). For each model, the fitness function has been calculated as the mean value between specificity and sensitivity obtained from the models built in each bootstrap iteration. This fitness function has been used to select the best descriptor to be added to proceed to the next iteration. From this procedure, the set of descriptors with the best cross-validation values has been chosen as the set to be used for the final model. The "best" values have been considered taking into account their trend: by progressively adding descriptors to the model, the cross-validation performances increase until a plateau (and, following, a decrease), which means that the optimal number of descriptors have been reached, and adding further descriptors would lead to over-fitting

4.5.Algorithm and descriptor generation:

VEGA platform, where the descriptors have been implemented following the definition available in the Dragon software

4.6.Software name and version for descriptor generation:

DRAGON

Dragon is an application for the calculation of molecular descriptors (MDs) and fingerprints (FPs).These can be used to evaluate molecular structure-activity or structure-property relationships, as well as for similarity analysis and high throughput screening of molecule databases. https://chm.kode-solutions.netKode srl, Dragon (software for molecular descriptor calculation) version 7.0.8, 2017, https://chm.kode-solutions.net

VEGA http://www.vegahub.eu (Emilio Benfenati, Alberto Manganaro)http://www.vegahub.eu

4.7.Chemicals/Descriptors ratio:

264/8 = 33

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 > 0.85, the predicted substance is regarded in the Applicability Domain of the model. It corresponds to “good reliability” of prediction.

If 0.85 ≥ 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 AD index < 0.6, 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.

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

Descriptors Range (*IdxDescRange*) index is calculated as 1.0 if all molecular descriptors used in the prediction fall within the range of descriptors used in the whole training set, 0.0 otherwise.

AD final index is calculated as following:

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

5.2. Method used to assess the applicability domain:

The 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 [4]. The VEGA 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 index < 0.6, 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 ≥ 0.8, 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 index < 0.6, 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 ≥ 0.8, similar molecules found in the training set have experimental values that completely disagree with the target compound predicted value

Model descriptors range 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, the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability

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 RARE * NOTFOUND. Defined intervals are:

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 atoms 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

Included in the VEGA software and automatically displayed when running the model

emilio.benfenati@marionegri.it

https://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.

Model is applicable for the Molecular weight between 18<MW<750 and octanol water partition coefficient between (-3 to +6)

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: null

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:

NA

6.6.Pre-processing of data before modelling:

The training set is composed of 264 mono constituent organic compounds, 180 Sensitizer and 84 non-sensitizer.

6.7.Statistics for goodness-of-fit:

Training set: n = 264; Accuracy = 0.80; Specificity = 0.79; Sensitivity = 0.81

TP: 145, TN: 66, FP:18, FN: 35

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

Chemical Name: No

Smiles: Yes

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:

All

7.5.Other information about the external validation set:

NA

7.6.Experimental design of test set:

The dataset were randomly split into training and test set with respectively the 80% and the 20% of the compounds. The test set is composed of 68 mono constituent organic compounds, 46 Sensitizer and 22 non-sensitizer.

7.7.Predictivity - Statistics obtained by external validation:

Test set: n = 68; Accuracy = 0.71; Specificity = 0.82; Sensitivity = 0.65TP: 30 TN:18, FP:4, FN: 16. It is noted that the specificity measure is based on only 22 non-sensitizers and hence somewhat uncertain

External validation set in AD: n = 79; Accuracy = 0.75; Specificity = 0.36; Sensitivity = 0.83 TP:54, TN:5, FP:9, FN:11

External validation set could be out of AD: n = 93; Accuracy = 0.55; Specificity = 0.52; Sensitivity = 0.56 TP:35, TN:16, FP:15, FN:27

External validation set out of AD: n = 203; Accuracy = 0.47; Specificity = 0.28; Sensitivity = 0.61 TP:72, TN: 24, FP:62, FN:45

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:

Some reasons of activity related to chemical moieties have been identified a posterior

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:

NA

[1] Chaudhry, Q., Piclin, N., Cotterill, J., Pintore, M., Price, N. R., Chrétien, J. R., & Roncaglioni, A. (2010). Global QSAR models of skin sensitisers for regulatory purposes. Chemistry Central Journal, 4(1), S5. https://doi.org/10.1186/1752-153X-4-S1-S5

[2] Asturiol, D., Casati, S. and Worth, A. (2016). Consensus of classification trees for skin sensitisation hazard prediction. Toxicol In Vitro 36, 197-209. doi:10.1016/j.tiv.2016.07.014

[3] OECD. (2002 (now replaced) & 2010). Test No. 429: Skin Sensitisation: Local Lymph Node Assay. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/environment/test-no-429-skin-sensitisation_9789264071100-en

[4] 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

[5] Selvestrel, G., Robino, F., Baderna, D., Manganelli, S., Asturiol, D., Manganaro, A., Russo, M. Z., Lavado, G., Toma, C., Roncaglioni, A., & Benfenati, E. (2021). Sphera Cosmolife: A new tool for the risk assessment of cosmetic products. ALTEX - Alternatives to Animal Experimentation, 38(4), 565–579. https://doi.org/10.14573/altex.2010221

[6] 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

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