	QMRF identifier (JRC Inventory): **Q17-46-0053**
	QMRF Title: **TIMES (TIssue MEtabolism Simulator) model for Skin sensitization (TIMES-SS model)**
	Printing Date: **Dec 11, 2019**

1.QSAR identifier

1.1.QSAR identifier (title):

TIMES (TIssue MEtabolism Simulator) model for Skin sensitization
(TIMES-SS model)

1.2.Other related models:

A module within the TIMES platform

1.3.Software coding the model:

Skin sensitization with autoxidation v.20.24
OASIS TIMES v.2.27.17
LMC, University "Prof. As. Zlatarov", Bourgas, Bulgaria e-mail: omekenya@btu.bg
http://www.oasis-lmc.org/

2.General information

2.1.Date of QMRF:

June, 2015

2.2.QMRF author(s) and contact details:

Ovanes Mekenyan LMC, University "Prof. As. Zlatarov", Bourgas, Bulgaria omekenya@btu.bg
http://www.oasis-lmc.org/

2.3.Date of QMRF update(s):

November 2013; June 2014; January 2015; June 2015

2.4.QMRF update(s):

Information which has been modified:
-Information in section 1. QSAR Identifier, field 1.3 Software coding
the model
-Information in Section 5. Defining the applicability domain of the
model – OECD Principle 3, fields 5.3 Software name and version for the
applicability domain assessment

2.5.Model developer(s) and contact details:

[1]Saby Dimitrov LMC Laboratory of Mathematical Chemistry using funding and data from a Consortium comprising Industry (ExxonMobil, P&G, Unilever, L'Oreal, Dow Chemicals, DuPont, Givaudan and RIFM) and a Regulatory Agency (DK-EPA). sdimitrov@btu.bg
[2]Ovanes Mekenyan LMC Laboratory of Mathematical Chemistry using funding and data from a Consortium comprising Industry (ExxonMobil, P&G, Unilever, L'Oreal, Dow Chemicals, DuPont, Givaudan and RIFM) and a Regulatory Agency (DK-EPA). omekenya@btu.bg
[3]Gergana Dimitrova LMC Laboratory of Mathematical Chemistry using funding and data from a Consortium comprising Industry (ExxonMobil, P&G, Unilever, L'Oreal, Dow Chemicals, DuPont, Givaudan and RIFM) and a Regulatory Agency (DK-EPA). geri_d@btu.bg

2.6.Date of model development and/or publication:

April 2005

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

Dimitrov S, Low L, Patlewicz G, Kern P, Dimitrova G, Comber M, Philips R, Niemela J, Bailey P, Mekenyan O (2005). Skin sensitization: modeling based on skin metabolism simulation and formation of protein conjugates. International Journal of Toxicology. 24, 189-204.

2.8.Availability of information about the model:

http://oasis-lmc.org/products/models/human-health-endpoints/skin-sensitization.aspx

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

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Mouse; guinea pigs

3.2.Endpoint:

4.Human Health Effects 4.6.Skin sensitisation

3.3.Comment on endpoint:

Semi quantitative potency score - strong, weak and non sensitising

3.4.Endpoint units:

LLNA – EC3, %

GPMT - % of animals showing reaction of skin

3.5.Dependent variable:

Obs. Skin Sensitization effect

3.6.Experimental protocol:

LLNA (the murine local lymph node assay); GPMT (the guinea pig maximization test)

3.7.Endpoint data quality and variability:

High quality. The model was derived from a data set compiled from chemicals tested in the LLNA, GPMT as well as from the BfR (formerly BgVV) list

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

QSAR

4.2.Explicit algorithm:

TIMES-SS model

TIMES-SS model aims to encode structure toxicity and structure metabolism relationships through a number of transformations simulating skin metabolism and interaction of the generated reactive metabolites with skin proteins. The skin metabolism simulator mimics metabolism using 2D structural information. The autoxidation (abiotic oxidation) of chemicals is also accounted for. A training set of diverse chemicals was compiled and their skin sensitization potency assigned to one of three classes. These three classes were Strong, Weak or Non sensitizing.

4.3.Descriptors in the model:

[1]EHOMO - Energy of the Highest Occupied Molecular Orbital [eV]

[2]ELUMO - Energy of the Lowest Unoccupied Molecular Orbital [eV]

[3]Molecular weight (MW)

[4]Electronegativity – 0.5*(EHOMO –ELUMO) [eV]

[5]E_GAP – (EHOMO –ELUMO) [eV]

[6]Log Kow

[7]ACCEPT_DLC – Acceptor superdelocalizability

4.4.Descriptor selection:

Descriptors were selected by using the probabilistic approach for identifying common stereoelectronic (reactivity) patterns of the chemicals – COREPA

4.5.Algorithm and descriptor generation:

The COREPA (COmmon REactivity PAttern) method [sect. 9.2, ref.3] was used to derive the sub-models incorporated in the TIMES-SS models. It is a probabilistic technique for identifying common stereoelectronic (reactivity) patterns of structurally diverse chemicals which may exert similar or differential biological effects. All energetically reasonable conformers are used to establish conformer distributions across the global and local stereoelectronic descriptors associated with the activity of studied chemicals. The COREPA model is derived in the form of a decision tree. Its logic boxes consist of decision rules based on the reactivity patterns described by a combination of global descriptors of molecular steric and electronic structure and local reactivity parameters associated with specific alerting groups

Two additional 3D QSAR models implemented into the TIMES-SS model were derived for predicting skin sensitization potential of Aldehydes and Michael acceptors:

1. 3D QSAR model distinguishing skin sensitizing from not skin sensitizing aldehydes:

-The chemicals used to derive the model were aldehydes;

-The decision tree is consisted of one node separating Strong from Weak and non-sensitizing aldehydes based on calculated EHOMO in the ranges [-11.7, -8.22] [-11.7, -11] and MW in the ranges [13.6, 184] [250, 254].

2. 3D QSAR model distinguishing skin sensitizing from not skin sensitizing Michael acceptors:

-The chemicals used to derive the model were Michael acceptor having a double bond adjacent to electron-withdrawing group.

The decision tree is consisted of two nodes. The first one separates Strong from Weak and Non-sensitizing chemicals based on calculated ELECTRONEGATIVITY in the ranges [-5.13, -4.23] [-6.02, -5.42] and E_GAP in the ranges [8.07, 10.5] [10, 10.9]. The second one separates Weak from Non-sensitizing Michael acceptors based on calculated Log(Kow) in the ranges [0.68, 6.53] [-1.61, -0.583] and ACCEPT_DLC in the ranges [0.229, 0.275] [0.233, 0.242].

4.6.Software name and version for descriptor generation:

COREPA-M software developed at Laboratory of Mathematical Chemistry

4.7.Chemicals/Descriptors ratio:

875/7=125

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

The applicability domain of TIMES-SS model consists of the following layers:

1. General parametric requirements - includes ranges of variation of log KOW and MW. It specifies in the domain only those chemicals that fall in the range of variation of the MW and log Kow defined on the bases of the correctly predicted training set chemicals.
This layer of the domain is applied only on parent chemicals.
2. Structural domain - it is represented by list of atom - centered fragments extracted from the chemicals in the training set. The training chemicals were split into two subsets: chemicals correctly predicted by the model and incorrectly predicted chemicals. These two subsets of chemicals were used to extract characteristics determining the "good" and "bad" space of the domain. Extracted characteristics were split into three categories: unique characteristics of correct and incorrect chemicals (presented only in one of the subsets) and fuzzy characteristics presented in both subsets of chemicals.
Structural domain is applied on parent chemicals, only.
3. Mechanistic domain - in SS model it includes:
-Interpolation space: this stage of the applicability domain of the model holds only for chemicals for which an additional COREPA model is required. It estimates the position of the target chemicals in the population density plot built in the parametric space defined by the explanatory variables of the model by making use the training set chemicals. Currently, the accepted threshold of population density is 10%.
The mechanistic domain is applied on the parent structures and on their metabolites.

5.2.Method used to assess the applicability domain:

A stepwise approach for determining the applicability domain of the TIMES-SS model is proposed, distinguishing chemicals for which the model provides highly reliable predictions.
General parametric requirements are imposed in the first stage, specifying in the domain only those chemicals that fall in the range of variation of the physicochemical properties of the chemicals in the training set. The second stage defines the structural similarity between chemicals that are correctly predicted by the model. The structural neighborhood of atom-centered fragments is used to determine this similarity. The third stage in defining the domain is based on a mechanistic understanding of the modeled phenomenon. Here, the model domain combines the reliability of specific reactive groups hypothesized to cause the effect and the domain of explanatory variables determining the parametric requirements in order for functional groups to elicit their reactivity.

5.3.Software name and version for applicability domain assessment:

Domain Manager v.1.09 developed at Laboratory of Mathematical Chemistry University, "Prof. Assen Zlatarov," 1 Yakimov Str., Bourgas 8010, BULGARIA

5.4.Limits of applicability:

In order to belong to the model applicability domain a target structure must meet the requirements of all the domain layers.

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

No

6.2.Available information for the training set:

CAS RN: Yes

Chemical Name: Yes

Smiles: Yes

Formula: Yes

INChI: No

MOL file: 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:

Training set consists of 875 chemicals (not attached).

<html><body>The current skin sensitization model was developed using a dataset of 875 chemicals tested by Local Lymph Node Assay (LLNA), Guinea Pig Maximization Test (GPMT) and chemicals from the BfR list.
A unifying scale was derived evaluating the correlation and concordance of those chemicals that existed in all three datasets:
Unified skin sensitization scale was proposed of - strong, weak, non - where strong corresponded to extreme, strong, moderate in the LLNA, strong & moderate in the GPMT and Category A in the BfR; weak corresponded to weak in the LLNA and GPMT and Category B in the BfR and Non was non in the LLNA, GPMT and Category C in the BfR.
The distribution of training set chemicals having skin sensitization experimental data among the sensitization classes is as follows:
-398 are Strong skin sensitizers
-193 are Weak skin sensitizers
-284 are Non skin sensitizers

6.6.Pre-processing of data before modelling:

6.7.Statistics for goodness-of-fit:

For 875 chemicals, the TIMES-SS model was able to predict correctly 91% of the strong sensitizers, 51% of the weak sensitizers and 69% of the non-sensitizers, i.e., an overall performance of 75 %.
Sensitivity: 77 %
Specificity: 69 %

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

Not provided

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: 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:

No

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

No

7.5.Other information about the external validation set:

External validation of the model was done by using a set of chemicals having skin sensitization data [sect. 9.2, ref.1].

7.6.Experimental design of test set:

7.7.Predictivity - Statistics obtained by external validation:

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:

The TIMES-SS (Tissue Metabolism Simulator for skin sensitization) model integrates a simulator of skin metabolism together with a number of “local” QSAR models for assessing the reactivity of specific alerts. A skin metabolism simulator was developed based on empirical and theoretical knowledge (not enough reported observed skin metabolism data). The transformation probabilities (defining the priority of their execution) were parameterized to reproduce skin sensitization data. The simulator comprises of about 420 transformations, which can be divided into four main types: abiotic transformations, covalent interaction with proteins, Phase I and Phase II reactions. Autoxidation (AU) of chemical is also accounted for. Interactions with skin proteins are grouped into three types: leading to strong or weak skin sensitization effect and interactions requiring QSAR models to quantify the potency of sensitization of the alerting groups. The QSAR models were developed by the COmmon PAttern Recognition (COREPA) approach [sect. 9.2, ref.3]. The skin sensitization model predicts skin sensitization effect in three classes: strong, weak and non-sensitizers.

Reliability of alerts in the TIMES-SS model has been also evaluated to

provide transparent mechanistic reasoning for predicting sensitization potential. Alert performance was defined as the ratio between the number of correct (positive and negative) predictions and the total number of chemicals within the local training set that triggered the alert. The alert performance was assessed based on the predictions on parents, autoxidation products simulated by the external AU simulator and metabolites as simulated by the skin metabolism simulator embedded in TIMES-SS model. Four different categories of reliability were defined:
-High reliability – alert performance higher than 60% and more than 5 chemical in local (transformation/alert) training set
-Low reliability – performance less than 60% and more than 5 chemicals in training set
-Undetermined reliability – less than 5 chemicals in training set
-Undetermined (theoretical) – there are no chemicals supporting the alert in the local training set</body></html>

8.2.A priori or a posteriori mechanistic interpretation:

8.3.Other information about the mechanistic interpretation:

9.Miscellaneous information

9.1.Comments:

The model can be used to predict skin sensitization potential of organic chemicals.

9.2.Bibliography:

[1]Dimitrov S, Low L, Patlewicz G, Kern P, Dimitrova G, Comber M, Philips R, Niemela J, Bailey P, Mekenyan O (2005). Skin sensitization: modeling based on skin metabolism simulation and formation of protein conjugates. International Journal of Toxicology. 24, 189-204.

[2]Dimitrov S, Low L, Patlewicz G, Kern P, Dimitrova G, Comber M, Aptula A, Philips R, Niemela J, Madsen C, Wedebye E, Robert D, Bailey P, Mekenyan O (2007). TIMES-SS - A promising tool for the assessment of skin sensitization hazard. A characterization with respect to the OECD validation principles for (Q)SARs and an external evaluation for predictivity, Regulator Toxicology and Pharmacology. 48, 225-23.

[3]Mekenyan O, Nikolova N, Schmieder P and Veith G (2004) COREPA-M, A multi-dimentional formulation of COREPA, QSAR & Combinatorial Science. 23 (1) 5-18.

9.3.Supporting information:

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

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

Q17-46-0053

10.2.Publication date:

2017-09-27

10.3.Keywords:

skin sensitisation;TIMES-SS;guinea pig maximization test;GPMT;Tissue Metabolism Simulator;Laboratory of Mathematical Chemistry;LMC;

10.4.Comments:

old# Q49-52-53-485