(Q)SAR Model Reporting Format (QMRF)

(The present QMRF is prepared according the fields and *Help* recommendations of JRC implemented in QMRF v 1.3 and QMRF Editor v.2.0.0 (https://sourceforge.net/apps/mediawiki/qmrf/index.php?title=Main_Page**)**

Welcome

Model version: Ames Mutagenicity S9 activated kinetic v.04.04
Platform version: OASIS TIMES 2.32.1
Name: Ames mutagenicity
Author: LMC, University "Prof. As. Zlatarov", Bourgas, Bulgaria
Date: 24 March 2023
E-mail: omekenya@btu.bg
www: http://www.oasis-lmc.org/

Section 1. QSAR identifier

1.1. QSAR identifier (title)

Ames Mutagenicity S9 activated kinetic model

1.2. Other related models

Not applicable

1.3. Software coding the model

Model version: Ames Mutagenicity S9 activated kinetic v.04.04
Platform version: OASIS TIMES 2.32.1
Name: *In vitro* Ames Mutagenicity S9 activated kinetic model
Developer: LMC, University "Prof. As. Zlatarov", Bourgas, Bulgaria
Coding language: Delphi 10.2

Section 2. Date of QMRF

2.1. Date of QMRF

24 March 2023

2.2. QMRF author(s) and contact details

Name: Laboratory of Mathematical Chemistry

Affiliation: Laboratory of Mathematical Chemistry, University "Prof. As. Zlatarov", "Yakimov" St. #1, 8010 Bourgas, BULGARIA

URL: http://www.oasis-lmc.org

E-mail: omekenya@btu.bg

2.3. Date of QMRF update(s)

29 November 2021, 24 March 2023

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 2. Date of QMRF,** fields **2.1 Date of QMRF; 2.3 Date of QMRF update(s); 2.8 Availability of information about the model**
- Information in **Section 3. Defining the endpoint – OECD Principle 1,** field **3.7 Endpoint data quality and variability**
- **Information in Section 4. Defining the algorithm – OECD Principle 2,** fields **4.4 Descriptor section; 4.6 Software name and version for descriptor generation**
- Information in **Section 5. Defining the applicability domain of the model – OECD Principle 3,** field **5.4. Limits of applicability**
- **Information in Section 6. Internal validation - OECD Principle 4,** fields **6.1 Availability of the training set; 6.7 Statistics for goodness-of-fit;**
- Information in **Section 7. External validation – OECD Principle 4,** fields **7.1 Availability of the external validation set; 7.2 Available information for the external validation set; 7.5 Other information about the external validation set; 7.7 Predictivity – Statistics obtained by external validation; 7.9 Comment on the external validation of the model**

2.5. Model developer(s) and contact details

Name: P. Petkov, A. Chapkanov, C. Kuseva, H. Ivanova, E. Kaloyanova, G. Dimitrova, D. Yordanova, R. Serafimova, M. Todorov, T. Pavlov, S. Kotov, E. Jacob, A. Aptula, O. Mekenyan

Affiliation: Laboratory of Mathematical Chemistry, University "Prof. As. Zlatarov", "Yakimov" St. #1, 8010 Bourgas, BULGARIA

URL: http://www.oasis-lmc.org

E-mail: omekenya@btu.bg

2.6. Date of model development and/or publication

2020

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

1. R. Serafimova, M. Todorov, T. Pavlov, S. Kotov, E. Jacob, A. Aptula, O. Mekenyan. 2007. Identification of the structural requirement for mutagenicity by incorporating molecular flexibility and metabolic activation of chemicals. II. General Ames mutagenicity model. *Chem. Res. Toxicol*., 662-676.
2. O. Mekenyan, S. Dimitrov, T. Pavlov, G. Dimitrova, M. Todorov, P. Petkov & S. Kotov. 2012. Simulation of chemical metabolism for fate and hazard assessment. V. Mammalian hazard assessment, *SAR and QSAR in Environmental Research*, Vol. 23, 553-606

2.8. Availability of information about the model

TIMES_Ames mutagenicity kinetic model (+S9) is derived for identification of chemicals capable to interact with DNA. Training set of the model includes 4196 chemicals (part of which are proprietary data) from different literature sources. The model is based on an alerting group approach addressing mutagenicity of parents and their generated metabolites *in vitro* liver S9 metabolic system taking into account the kinetics of metabolism. The model allows to relate the level of mutagenic potency to the amount of formed DNA adducts under the assumption that the presence of alert is necessary but not sufficient reason for predicting positive mutagenicity.

2.9. Availability of another QMRF for exactly the same model

Not available.

Section 3. Defining the endpoint – OECD Principle 1

3.1. Species

Chemicals included in the training set of the TIMES_Ames kinetic model are collected according to the recommendation in the OECD technical guideline 471 addressing the number of *Salmonella typhimurium* strains (TA100, TA98, TA1535, TA1537, *E. coli*) associated with each data:

- For negative effect, all five *Salmonella* strains must show simultaneously negative data as described in the corresponding OECD guideline for testing of chemicals:

http://www.oecd-ilibrary.org/environment/test-no-471-bacterial-reverse-mutation-test_9789264071247-en.

- For positive effect, positive data in a single *Salmonella* strain would be enough.

3.2. Endpoint

Bacterial Reverse Mutation Test

According to JRC pre-classification list of endpoints:
No. 207 QMRF Human Health Effects, QMRF 4.10 Mutagenicity.

3.3. Comment on endpoint

To detect and measure potency of DNA mutagens, a short-term, simple and inexpensive *in vitro* assay is used, such as the Bacterial reverse mutation (Ames) test. The Ames test detects single nucleotide base change, base insertion or deletion in different *Salmonella* strains. All *Salmonella* strains carry some type of defective (mutant) gene that prevents them from synthesizing the amino acid histidine. In the presence of mutagenic chemicals, the defective gene may be mutated back to the functional state allowing the bacterium to grow.

3.4. Endpoint units

Qualitative – positive/ negative

3.5. Dependent variable

Observed Mutagenicity with S9

3.6. Experimental protocol

OECD technical guideline 471: Bacterial Reverse Mutation Assay (e.g. Ames test).

3.7. Endpoint data quality and variability

References associated with each documented mutagenicity data (except for proprietary data) included in the training set of the model are provided in ***Appendix 1***.

Section 4. Defining the algorithm – OECD Principle 2

4.1. Type of model

Structural alerts based model

4.2. Explicit algorithm

Prediction of Bacterial (Ames) mutagenicity is based on modelling of the two events deemed to be crucial for the effect – interaction of the chemicals with DNA and their activation as a result of liver S9 metabolism.

In the new modelling concept the presence of alerts is necessary but not sufficient reason for predicting a positive effect. It requires taking into account the kinetics of metabolism. In the new (kinetic) in vitro S9 metabolic simulators:

- Experimental kinetic data (clearances) are used to calculate the probability of transformations as a function of time already, i.e. P = (1-exp(-Cl*t)).
- Clearance data have been used to optimize the probability of metabolic transformations.

  In addition, expert information for the stability of chemicals in rat S9 mix is also used.

As a result, the transformations in the metabolic simulator are modified to simulate the formation of DNA adducts (not existing in the original models). Mutagenic potency is associated with concentration of the formed DNA adducts which is estimated over time. The magnitude of these adducts is assumed to correspond to the level of damage of macromolecules and potency effect, respectively. As a result empirical thresholds of the formed DNA adducts are derived to distinguish positive from negative chemicals.

Details about the alerts included in the model are provided in the next sections.

4.3. Descriptors in the model

Descriptors in the model are structural alerts related to interactions with DNA. Alerts in the TIMES Ames kinetic model (+S9) constitute expertly-derived sets of structural fragments incorporating knowledge for the interactions of chemicals (parents and metabolites) with DNA.

Description of these alerts is provided in the next sections.

4.4. Descriptor section

The main characteristics of each DNA alert in TIMES_Ames (+S9) kinetic model are as follows:

- Alert name (corresponding to the name of the chemical class which is addressed);
- Performance of alert (correct/incorrect predictions) which is estimated based on proportion of observed positive chemicals from all chemicals captured by the alert. Performance of each alert is provided with its confidence range. As smaller is the size of local training sets as wider are the confidence ranges and vice versa.
- P-values addressing the reliability of alert performance estimation and taking into account possible bias of positive/negative chemicals in the training set of the model. Low p-values could be obtained only if both are satisfied:
    - The number of chemicals in local training set is high enough;
    - The alert performance is significantly higher than the proportion of positive/negative chemicals in the model training set, i.e. so-called naïve alert.

Analogically, high p-values could be obtained in case of:

- Small number of local training set chemicals (1-2 chemicals); or
- Performance comparable to the performance of the naïve alert.

High performance associated with low *p-values* indicates for High Reliability of alerts. Full list of the alerts in TIMES_Ames (+S9) kinetic model with defined quantity thresholds is given in Table 1:

Table 1. List with DNA alerts in the TIMES Ames kinetic model (+S9).

No	Alert Name	Defined Quantity threshold for	
		parents	metabolites
1	1,4-Diazabutadiene Derivatives		
2	4,4'-Bipyridinium Salts and N-Oxides		
3	Acridone, Thioxanthone, Xanthone, Phenazine and Other Fused-Ring Heterocyclic DNA Intercalators		0.09
4	Acyclic Triazenes		
5	Acyl Halides		

6	Acyl Halides		
7	Aliphatic saturated monoaldehydes		
8	Alkyl Xanthate Esters		
9	Alkylated nitrosoureas and nitrosoguanidines		
10	Alkylnitrites		
11	Alkylphosphates, Alkylthiophosphates and Alkylphosphonates	0.51	0.25
12	Alpha,Beta-Unsaturated Aldehydes		
13	alpha,beta-Unsaturated Carbonyls and Related Compounds		
14	alpha,beta-Unsaturated Carboxylic Acids and Esters		
15	alpha,omega-Dihaloalkanes		
16	alpha-Activated benzyls		
17	alpha-Activated Haloalkanes		
18	Alpha-Beta Conjugated Alkene Derivatives with Geminal Electron-Withdrawing Groups		
19	Alpha-Haloethers		
20	Amidoxime Esters and Amidoximes		
21	Amino Anthraquinones		0.125
22	Aminoacridine DNA Intercalators		
23	Amphetamine derivatives		
24	Anthrones		0.73
25	Antibiotic Aminoglycoside Derivatives		
26	Arenecarboxylic Acid Esters		
27	Arenediazonium and Diazonium Salts		0.03
28	Arenesulfonamides		
29	Aromatic ester hydroxylamine		
30	Atrazine derivatives		
31	Azoalkanes with Activating Electron-Withdrawing Groups		
32	Azodicarbonamides		
33	Azoxyalkanes		0.038
34	Benzanthrone Derivatives		
35	Benzoquinoline and Acridine derivatives		
36	Bipyridilium Herbicides		
37	Bleomycin and Structurally Related Compounds		

38	Carbamates		
39	Carboxylic Acid Amides		
40	Carboxylic Acid Anhydrides		
41	Chlorinated Diphenylmethane and Benzophenone Derivatives		
42	C-Nitroso Compounds		
43	Conjugated Benzoylethene Derivatives		
44	Conjugated Nitroalkenes and Five-Membered Aromatic Nitro- and Amino Heterocycles		
45	Coumarins and Thiocoumarins		
46	Coumarins and Thiocoumarins		
47	Coumarins and Thiocoumarins		
48	Cyanohydrins		
49	Cyclic maleic acid derivatives		
50	Dialkyl Alkylphosphonates		
51	Diazenes		0.021
52	Diazoalkanes		
53	Dicarbonyl compounds		0.35
54	Dichlorophosphine and Dichlorophosphonium Derivatives		
55	Dithianes		
56	DNA Intercalators with Carboxamide and Aminoalkylamine Side Chain		
57	Epoxides, Aziridines, Thiiranes and Oxetanes	0.53	0.128
58	Ethenyl Pyridines		
59	Flavonoids		
60	Formaldehyde Releasers		
61	Four- and Five-Membered Lactones		
62	Fused-Ring Conjugated Lactones		
63	Fused-Ring Nitroaromatics		
64	Fused-Ring Primary Aromatic Amines		
65	Gallic Acid Esters		
66	Geminal Polyhaloalkane Derivatives		
67	Haloalcohols		
68	Haloalkane Derivatives Containing Chain Heteroatom		

69	Haloalkane Derivatives with Labile Halogen		0.04
70	Haloalkene Cysteine S-Conjugates		
71	Haloalkene Derivatives with Electron-Withdrawing Groups		
72	Haloazaarene and Fused-Ring Haloquinoline Derivatives		
73	Halofuranones		
74	Halogenated Oxetanes and Haloepoxides		
75	Halogenated Vicinal Hydrocarbons		
76	Haloisothiazolinones		
77	Heteroarene Sulfenamides		
78	Heterocyclic Aromatic Amines		
79	Heterocyclic N-Hydroxylamines		0.03
80	Heterocyclic nitro compounds		
81	Heterocyclic Nitroso compounds		
82	Heterocyclic urea derevatives		
83	Hexahydrotriazine Derivatives		
84	Hydrazine Derivatives	0.37	
85	Hydroxamic Acids		0.31
86	Hydroxybenzophenone Derivatives		0.06
87	Hydroxylated Phenols		
88	Hypoxanthine Derivatives		
89	Imidazolinone derivatives		
90	Isocyanates and Diisocyanates		
91	Lactones		
92	Monohaloalkanes		
93	N,N-Dialkyldithiocarbamate Derivatives and Azaarene Dithiocarbamates		
94	N-acetoxyamines		
95	N-Acyloxy(Alkoxy) Arenamides		
96	N-Alkylindolinium and N-Alkylbenzothiazolium Salts		
97	N-Alkyl-N-nitrosocarbamates		
98	N-Aryl-N-Acetoxy(Benzoyloxy) Acetamides		
99	N-Hydroxyethyl Lactams		
100	N-Hydroxylamines	0.51	0.062
101	Nitroalkanes – Mononitroalkanes		

102	Nitroaniline Derivatives		0.3
103	Nitroarenes with Other Active Groups		0.18
104	Nitroazoarenes and p-Monosubstituted Azobenzene Derivatives		
105	Nitrobiphenyls and Bridged Nitrobiphenyls		
106	Nitrogen and Sulfur Mustards		
107	Nitrogen mustards		
108	Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids	0.6	0.1
109	N-methylol derivatives		
110	N-Nitroso Compounds	0.55	0.08
111	N-Nitrosoamine derivatives		
112	Non-aromatic conjugated systems with electron-withdrawing groups		
113	Non-Aromatic Hydroxylamine Derivatives		
114	Non-Cyclic Alkyl Phosphoramides and Thionophosphoramides		
115	N-Substituted Aromatic Amines		
116	N-Trihalomethyl Imides		
117	Organic Azides		
118	Organic Diselenides and Ditellurides		
119	Organic Peroxy Compounds		
120	PAH Benzylic Alcohol Esters		
121	p-Aminobiphenyl Analogs		
122	Perfluorinated Hypofluorites		
123	Peroxyacyl Nitrates		
124	Polarized Haloalkene Derivatives		
125	Polycyclic Aromatic Hydrocarbon, Naphthaleneimide and Carbazole Derivatives		
126	Polyethylene Polyamines		
127	Polynitroarenes		0.5
128	Propargyl Alcohol Derivatives		
129	Propyne Derivatives		
130	p-Substituted Mononitrobenzenes		
131	Pyrazolone and Pyrazolidine-3,5-dione Derivatives		
132	Pyrrolizidine Derivatives		

133	Isothiocyanates		
134	Quinoline Derivatives	0.5	
135	Quinolone Derivatives		
136	Quinone methides		
137	Quinoneimine, Thionine and Phenoxazinium Derivatives		0.05
138	Quinoneimines protein binding		
139	Quinones and Trihydroxybenzenes		0.02
140	Quinoxaline-Type 1,4-Dioxides		
141	S-Activated Cysteine Derivatives		
142	Short-Chain Alkyltin and Alkylgermanium Halides		
143	Single-Ring Substituted Primary Aromatic Amines	0.15	0.24
144	Specific 5-Substituted Uracil Derivatives		
145	Specific Acetate Esters		
146	Specific Imine and Thione Derivatives	0.47	0.4
147	Sterically Hindered Piperidine Derivatives		
148	Substituted Anilines		
149	Substituted Benzoindoline and Indole Derivatives		
150	Substituted Chlorophenylalkylurea Derivatives		
151	Substituted Nitropyridines, Aminopyridines and N-Oxides		
152	Substituted Phenols		
153	Sulfonates and Sulfates	0.71	0.045
154	Sulfonyl Azides		
155	Sulfonyl Halides		
156	Sultones		
157	Tertiary aromatic amine		
158	Thiadiazole-dioxide derivatives		
159	Thiazolidinediones		
160	Thiocarbonyl S,S-dioxides		
161	Thiols	0.885	0.45
162	Triarylimidazole and Structurally Related DNA Intercalators		
163	Tri-Methylindole derivatives		
164	Vicinal Dihaloalkanes		

4.5. Algorithm and descriptor generation

The structural boundaries of the alerts are derived from the chemicals included in the local training sets. For derivation of each alert mechanistically justifiable structural fragments for interaction with DNA are identified from the chemicals having positive data in the local training set. Additional structural fragments from the other parts of the molecules which could affect (enhance or reduce) the mutagenicity effect are also introduced to complete definition of most alerts.

4.6. Software name and version for descriptor generation

Ames Mutagenicity S9 activated kinetic model version 04.04

4.7. Chemicals/Descriptors ratio

Not applicable

Section 5. Defining the applicability domain of the model – OECD Principle 3

5.1. Description of the applicability domain of the model

The domain consists of the following sub-domain layers:

1. General parametric requirements.

The variations of molecular parameters that may affect the quality of the measured endpoint significantly are included here (such as molecular weight, etc.). The domain of general parametric includes the range of variation of hydrophobicity (log K_{OW}) and Molecular weight (MW) of chemicals in training set.

2. Structural domain.

The structural component of the model is based on the structural similarity between chemicals in the training set which were correctly predicted by the model. The structural neighborhood of atom-centered fragments (accounting for the first neighbours) extracted from correctly and incorrectly predicted parent structures from the training set is used to determine this similarity.

The target chemical could contain the following types of ACF:

- Fragments present in correctly predicted training chemicals only (i.e. correct fragments),

- Fragments found both in correctly and non-correctly predicted training chemicals (i.e. fuzzy fragments). These fragments are treated as correct fragments,
- Fragments present in non-correctly predicted training chemicals only (i.e. incorrect fragments),
- Fragments not present in the training chemicals (i.e. unknown fragments).

A chemical belongs to the structural domain of the model if it could be partitioned only on correct fragments. The user is able to analyse how important are unknown and incorrect fragments (if present in the target) and to make a decision about their effect on the quality of prediction. The distribution of structural characteristics of the target chemical and accepted thresholds is used as a criterion to determine how well the target is represented in the structural space of correctly predicted chemicals. The accepted domain thresholds for Mutagenicity are as follows:

- Correct = 100%
- Incorrect = 0%

A chemical is considered In Domain if it is classified to belong to all sub-domain levels. The information implemented in the applicability domain is extracted from the correctly predicted training chemicals used to build the model and in this respect the applicability domain determines practically the interpolation space of the model.

5.2. Method used to assess the applicability domain

The approach used to determine and assess the domain is described in:

Dimitrov S, Dimitrova G., Pavlov T., Dimitrova N., Patlewicz G., Niemela J., Mekenyan O., A stepwise approach for defining the applicability domain of SAR and QSAR models, *J. Chem. Inf. Model.*, 45, 839-849 (2005).

5.3. Software name and version for the applicability domain assessment

The LMC software OASIS Domain Manager v.1.13 (which is embedded in OASIS platform) is used to determine the applicability domain.

http://oasis-lmc.org/products/software/domain-manager.aspx

5.4. Limits of applicability

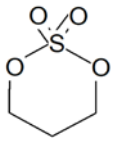
Applicability domain of the Ames kinetic model (+S9) includes three sub-domain layers: general parametric requirements, structural features and alerts reliability.

- General properties requirements:

As described in the Section 5.1.1, parametric domain of the model is derived based on Log *Kow* and *MW*. Example demonstrating belonging of a training set chemical to the parametric layer of the model domain is provided below:

Example chemical:

- CAS: 1073-05-8
- Name: 1,3, 1,3,2-dioxathiane 2,2-dioxide
- 2D Depiction:



Property	Domain	Example chemical
log K_{OW}	[-17.64; 24.665]	-0.410
MW, Da	[31.025; 1561.757]	138.137

* K_{OW} is calculated by EPIWin

The values of *logKow* and *MW* of the example chemical are within the ranges of these parameters extracted from the whole training set of the model. Hence, with respect to the general parametric requirements, the example chemical is estimated to be *In Domian*.

- Structural features

  Structural domain of the model is extracted from 4196 training chemicals contains:
  - 9812 correct fragments,
  - 964 fuzzy fragments (treated as correct fragments),
  - 800 incorrect fragments.

- Alerts reliability

Reliability of alerts is estimated based on:

- Alert performance of the local training set chemicals (AP);
- Number of the local training sets (N);
- Mechanistic justification (M).

According to these criteria, there are four reliability estimates for the alerts in the models:

- High reliability alerts (AP>0.6, N>10, M);
- Low reliability alerts (AP<0.6, N>10, M);
- Undetermined alerts (N<10, M);
- Undetermined theoretical alerts (M).

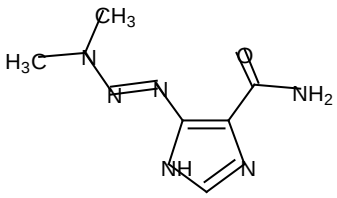
Example chemical belonging to alert with “High reliability”.

Chemical ID:

CAS: 4342-03-4

Name: Dacarbazine

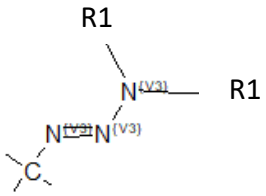
Depiction:



Belonging to alert:

Name: Acyclic Triazenes

Structural boundaries:



R_1 = H; CH_3; C_2H_5; $CH(CH_3)_2$; CH_2-C_6H_5

Reliability:

“High reliability” based on AP=1; N=15 and M.

Currently, information for alerts reliability is provided in the model reports.

Section 6. Internal validation – OECD Principle 4

6.1. Availability of the training set

Training set of the TIMES Ames kinetic model (+S9) includes 4196 organic compounds from different chemical classes.

6.2. Available information for the training set

CAS numbers, Chemical names, SMILES, documented data, literature sources and strain information are available for each compound in the model training set.

6.3. Data for each descriptor variable for the training set

Not applicable

6.4. Data for the dependent variable for the training set

Data for the dependent variable of the training set are embedded in the software implementation of the model.

6.5. Other information about the training set

Not available

6.6. Pre-processing of data before modelling

Not available

6.7. Statistics for goodness-of-fit

Statistics of the model:

- Sensitivity = (predicted positive/observed positive) = 89%
- Specificity = (predicted negative/observed negative) = 95%
- Concordance = (correct predicted positive and negative chemicals in respect to all training set chemicals) = 92%

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

Not performed

6.9. Robustness – Statistics obtained by leave-many-out cross-validation

Not performed

6.10. Robustness - Statistics obtained by Y-scrambling

Not performed

6.11. Robustness - Statistics obtained by bootstrap

Not performed

6.12. Robustness - Statistics obtained by other methods

Not performed

6.13. Comment on the internal validation of the model

Not performed

Section 7. External validation – OECD Principle 4

7.1. Availability of the external validation set

10 199 external chemicals are available to examine performance of the model.

7.2. Available information for the external validation set

According to the OECD TG 471, the external validation set addresses the five *Salmonella* strains (TA100, TA90, TA1535, TA1537 and *E. coli* WP2 uvrA).

7.3. Data for each descriptor variable for the external validation set

Not available

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

Not available

7.5. Other information about the external validation set

The list with 10 199 chemicals with Ames negative experimental data (Class C) are provided by the Division of Genetics and Mutagenesis of National Institute of Health Sciences of Japan. Details for the data used in the current external validation are available in the corresponding publication:

M. Honma, A. Kitazawa, A. Cayley, R. Williams, C. Barber, T. Hanser, R. Saiakhov, S. Chakravarti, G. Myatt, K. Cross, E. Benfenati, G. Raitano, O. Mekenyan, P. Petkov, C. Bossa, R. Benigni, C. Battistelli, A. Giuliani, O. Tcheremenskaia, C. DeMeo, U. Norinder, H. Koga, C. Jose, N. Jeliazkova, N. Kochev, V. Paskaleva, C. Yang, P. Daga, R. Clark, J. Rathman. 2019. Improvement of quantitative structure-activity relationship (QSAR) tools for predicting Ames mutagenicity: outcomes of Ames/QSAR International Challenge Project. Mutagenesis, Vol. 34, pp. 3-16.

7.6. Experimental design of test set

The external validation set contains only AMES negative substances.

7.7. Predictivity – Statistics obtained by external validation

- 8 979 out of 10 199 are predicted as Negatives by TIMES *in vitro* Ames kinetic model and these predictions are consistent with the experimental data, thus, the Specificity of the external set is 88% (8979/10199)
- 1 220 false positives
- Applicability domain - 1619 out of 10199 chemicals are *In domain*

7.8. Predictivity – Assessment of the external validation set

Not available

7.9. Comment on the external validation of the model

Performance of the TIMES_Ames kinetic model (+S9) with respect to non-mutagenic chemicals is quite high: the Specificity of the external set is 88% (8979/10199).

Section 8. Providing a mechanistic interpretation – OECD Principle 5

8.1. Mechanistic basis of the model

Only alerts extracted from the local training sets having clear interpretation of the molecular mechanism causing the mutagenicity effect are included in the model. Mechanistic rational of each alert is provided by experts based on significant reference support from the literature.

8.2. *A priori* or *a posteriori* mechanistic interpretation

The model building followed the traditional approach:

a. Building a hypothesis for the modelled event,
b. Defining the alerting groups based on parent structures,
c. Defining empirically the appropriate adduct quantity thresholds,
d. Fitting of model variable to the observed data,
e. Verification of model quality,
f. Depending on the results found in step *e* model building could continue with step *a, b, c* or *g,*
g. Determination of the applicability domain and practical application of the model.

8.3. Other information about the mechanistic interpretation

Additional information about the mechanistic interpretation could be fund in Section 2 (2.7).

Section 9. Miscellaneous information

9.1. Comments

Model predictions are fully transparent. The user is able to analyse the whole prediction process and to verify whether it concise with his/her knowledge or purposes.
For other related models, see Section 1 (1.2).

9.2. Bibliography

Additional references are not provided.

9.3. Supporting information

Additional supporting information is not provided.