(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: *In vitro* Chromosomal Aberrations v.17.17
Platform version: OASIS TIMES 2.30.1
Name: *In vitro* Chromosomal Aberrations
Author: LMC, University "Prof. As. Zlatarov", Bourgas, Bulgaria
Date: 22 Jan 2020
E-mail: omekenya@btu.bg
www: http://www.oasis-lmc.org/

Section 1. QSAR identifier

1.1. QSAR identifier (title)

In vitro Chromosomal Aberrations with S9 metabolic activation

1.2. Other related models

In vitro Ames mutagenicity with S9 metabolic activation

1.3. Software coding of the model

Model version: *In vitro* Chromosomal Aberrations v.17.17
Platform version: OASIS TIMES 2.30.1
Name: *In vitro* Chromosomal Aberrations
Developer: LMC, University "Prof. As. Zlatarov", Bourgas, Bulgaria
Coding language: Delphi 10.2

Section 2. Date of QMRF

2.1. Date of QMRF

22 Jan 2020

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)

21 November 2014; 12 June 2015; 12 May 2016; 12 July 2016; 31 August 2016; 30 May 2017; 23 July 2018; 22 Jan 2020

2.4. QMRF update(s)

Information which has been modified:
Section 2.7. Reference(s) to the main scientific and/or software package; **Sections 2.8.** Availability of information about the model; **Section 3.1.** Species; **Section 3.3.** Comments on endpoint; **Section 3.6.** Experimental protocol; **Section 4.2.** Explicit algorithm; **Section 4.3** Descriptors in the model; **Section 4.4.** Descriptor section; **Section 4.5.** Algorithm and descriptor generation; **Section 4.6.** Software name and version for descriptor generation; **Section 5.4.** Limits of applicability; **Section 6.** Internal validation – OECD Principle 4; **Section 7.** External validation – OECD Principle 4; **Section 8.1.** Mechanistic basis 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

2007/2012

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

1. O. Mekenyan; M. Todorov; R. Serafimova; S. Stoeva; A. Aynur; R. Finking; E. Jacob. Identifying the structural requirements for chromosomal aberration by incorporating molecular flexibility and metabolic activation of chemicals. Chem Res Toxicol, 1927-1941, (2007).
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_CA mutagenicity model (+S9) is derived for identification of chemicals capable to interact with DNA and proteins. Training set of the model includes 855 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. Details of the model is provided in the sections bellow as well as in the following link:

http://oasis-lmc.org/products/models/human-health-endpoints/chromosomal-aberrations.aspx

2.9. Availability of another QMRF for exactly the same model

Section 3. Defining the endpoint – OECD Principle 1

3.1. Species

Chemicals included in the training set of the TIMES_CA model are collected according to the recommendation in the OECD technical guideline 473 addressing: Chinese Hamster Ovary (CHO), Chinese Hamster lung V79 and Chinese Hamster Lung (CHL)/IU, TK6:

http://www.oecd-ilibrary.org/environment/test-no-473-in-vitro-mammalian-chromosomal-aberration-test_9789264224223-en

3.2. Endpoint

In vitro Mammalian Chromosome Aberration Test

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

3.3. Comment on endpoint

The purpose of the *in vitro* chromosomal aberration (CA) test is to identify substances that cause structural chromosomal aberrations in cultured mammalian cells. Structural aberrations may be of two types, chromosome or chromatid which are detected by the *in vitro* CA test. Polyploidy could arise in chromosome aberration assays *in vitro*. While aneugens can induce polyploidy, polyploidy alone does not indicate aneugenic potential and can simply indicate cell cycle perturbation or cytotoxicity. This test is not designed to measure aneuploidy.

3.4. Endpoint units

Qualitative – positive/ negative

3.5. Dependent variable

Obs. chromosomal aberrations with S9

3.6. Experimental protocol

OECD technical guideline 473: *in vitro* mammalian chromosomal aberration 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 alert based model

4.2. Explicit algorithm

Prediction of Chromosomal aberrations is based on modelling of the two events deemed to be crucial for the effect – interaction of the chemicals with DNA/proteins and their activation as a result of liver S9 metabolism.

CA mutagenicity predictions are obtained using an alerting group approach. Only alerts having clear interpretation of mechanisms leading to DNA mutagenicity are included in the model. To obtain predictions, a set of alerts (137) is applied on parents and each of the generated *in vitro* rat liver S9 metabolites. The *in vitro* S9 metabolic simulator is trained to reproduce documented maps for mammalian liver metabolism for 261 chemicals. Match of alerts either on parents or metabolites is sufficient for obtaining positive prediction. Chemicals are predicted to be mutagenic as parents only, parents and metabolites, or as metabolites only.

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 boundaries associated with alerting groups related to interactions with DNA/proteins. The definitions of the structural boundaries are derived based on subsets of chemicals from the whole training set having same alerting functionalities, i.e. "local" training set. Each alert is supported by detailed mechanistic description of the interaction with DNA/proteins provided as additional explanatory textual information. The alerts in the model are described with two main definitions:

- General mechanistic description provides broader characterization of the alerts. This definition is used only for defining the local training sets associated with specific alert. The procedure for obtaining local training sets includes searching of the general fragments among all chemicals from the training set of the model after application of S9 metabolic simulator. According to this, local training sets contain parent chemicals in which general fragments are:
    - found in their structures;
    - not found in the parent structures but found in their metabolite(s).

  General mechanistic alerts are not explicit for users because they are not intended for making predictions but just for derivation of local training sets.
- Endpoint-specific description provides more definitive characterization of the alerts which is developed for making predictions. Definitions of these alerts are more specific than the general ones accounting for the influence of other atoms or chemical groups found in the molecules. Description of these alerts is provided in the next sections.

4.4. Descriptor section

Table 1 summarizes the main characteristics of each DNA alert in TIMES_CA (+S9):

- 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* indicate for High Reliability of alerts.

The above statistical measures along with the underlying mathematical formalisms are discussed in details in **Section 6** (Internal validation).

Table 1. Main characteristics of the DNA alerts in the TIMES_CA model (+S9).

No.	**Alert**	**Correct**	**Incorrect**	**Performance $_{est.}$[1)]**	***p-value*[1)]**
1	N-Nitroso Compounds	41	1	0.955	1.4×10^{-7}

				(0.894 ÷ 0.999)	
2	Alkylated nitrosoureas and nitrosoguanidines	18	0	0.950 (0.854 ÷ 1.000)	0.0003
3	N-Alkyl-N-nitrosocarbamates	15	0	0.941 (0.829 ÷ 1.000)	0.0010
4	Quinoid compounds	45	2	0.939 (0.872 ÷ 0.993)	2.3×10^{-7}
5	Fused-Ring Primary Aromatic Amines	13	0	0.933 (0.807 ÷ 1.000)	0.0025
6	alpha,beta-Unsaturated Carboxylic Acids and Esters	26	1	0.931 (0.841 ÷ 0.998)	0.0001
7	Heterocyclic Aromatic Amines	12	0	0.929 (0.794 ÷ 1.000)	0.0039
8	Alkylphosphates, Alkylthiophosphates and Alkylphosphonates	10	0	0.917 (0.762 ÷ 1.000)	0.0099
9	Arenesulphonamides	10	0	0.917 (0.762 ÷ 1.000)	0.0099
10	Vicinal Dihaloalkanes	10	0	0.917 (0.762 ÷ 1.000)	0.0099
11	N-Substituted Aromatic Amines	19	1	0.909 (0.792 ÷ 0.997)	0.0013
12	Haloalkane Derivatives Containing Chain Heteroatom	17	1	0.900 (0.772 ÷ 0.997)	0.0029
13	Carbamates	8	0	0.900 (0.717 ÷ 1.000)	0.0247
14	Polarized Haloalkene Derivatives	8	0	0.900 (0.717 ÷ 1.000)	0.0247
15	Hydroxylated phenols	39	4	0.889 (0.797 ÷ 0.970)	0.0001
16	Quinoneimines protein binding	15	1	0.889 (0.748 ÷ 0.997)	0.0065
17	Nitrogen Mustards	7	0	0.889 (0.688 ÷ 1.000)	0.0391
18	N-nitrosoamine Derivatives	7	0	0.889 (0.688 ÷ 1.000)	0.0391
19	Polycyclic Aromatic Hydrocarbon and Naphthalenediimide Derivatives	7	0	0.889 (0.688 ÷ 1.000)	0.0391
20	N-Hydroxylamines	67	8	0.883	5.6×10^{-7}

				(0.811 ÷ 0.950)	
21	Substituted Anilines	50	6	0.879 (0.795 ÷ 0.956)	1.4×10^{-5}
22	C-Nitroso compounds protein binding	42	5	0.878 (0.786 ÷ 0.960)	0.0001
23	Halogenated Vicinal Hydrocarbons	13	1	0.875 (0.719 ÷ 0.996)	0.0144
24	Arenecarbonyl Compounds	6	0	0.875 (0.652 ÷ 1.000)	0.0620
25	Arenecarboxylic Acid Esters	6	0	0.875 (0.652 ÷ 1.000)	0.0620
26	Sulfonates and Sulfates DNA binding	6	0	0.875 (0.652 ÷ 1.000)	0.0620
27	Substituted Phenols	51	7	0.867 (0.780 ÷ 0.946)	3.1×10^{-5}
28	Pyrimidines and Purines	12	1	0.867 (0.701 ÷ 0.996)	0.0213
29	alpha,beta-Unsaturated Carbonyls and Related Compounds	24	3	0.862 (0.738 ÷ 0.971)	0.0030
30	C-Nitroso Compounds	36	5	0.860 (0.757 ÷ 0.954)	0.0005
31	Conjugated Nitroalkenes and Five-Membered Aromatic Nitroheterocyclics	5	0	0.857 (0.607 ÷ 1.000)	0.0984
32	Dialkyl Alkylphosphonates	5	0	0.857 (0.607 ÷ 1.000)	0.0984
33	Sulfonates and Sulfates protein binding	5	0	0.857 (0.607 ÷ 1.000)	0.0984
34	Haloalkane Derivatives with Labile Halogen	28	4	0.853 (0.735 ÷ 0.959)	0.0023
35	Epoxides, Aziridines and Thiiranes	54	9	0.846 (0.758 ÷ 0.928)	0.0001
36	(Thio)Phosphates	10	1	0.846 (0.659 ÷ 0.994)	0.0461
37	Epoxides and Aziridines	44	8	0.833 (0.734 ÷ 0.926)	0.0007
38	Benzoquinolines and Acridines derivatives	9	1	0.833 (0.632 ÷ 0.994)	0.0672
39	Fused-Ring Nitroaromatics	4	0	0.833 (0.549 ÷ 1.000)	0.1563

40	Isothiocyanates	4	0	0.833 (0.549 ÷ 1.000)	0.1563
41	Nitrogen and Sulfur Mustards	4	0	0.833 (0.549 ÷ 1.000)	0.1563
42	Pyrazolone and Pyrazolidine Derivatives	4	0	0.833 (0.549 ÷ 1.000)	0.1563
43	Single-ring Substituted Primary Aromatic Amines	44	9	0.818 (0.716 ÷ 0.914)	0.0015
44	Quinones and Trihydroxybenzenes	11	2	0.800 (0.605 ÷ 0.971)	0.0869
45	Quinoneimine, Thionine and Phenoxazinium Derivatives	7	1	0.800 (0.567 ÷ 0.991)	0.1401
46	Acyl Halides	3	0	0.800 (0.473 ÷ 1.000)	0.2483
47	alpha-Activated benzyls	3	0	0.800 (0.473 ÷ 1.000)	0.2483
48	Carboxylic Acid Anhydrides	3	0	0.800 (0.473 ÷ 1.000)	0.2483
49	Gallic Acid Esters	3	0	0.800 (0.473 ÷ 1.000)	0.2483
50	Quinone Methides	3	0	0.800 (0.473 ÷ 1.000)	0.2483
51	alpha-Activated Haloalkanes	28	7	0.784 (0.651 ÷ 0.907)	0.0255
52	Haloalcohols	17	4	0.783 (0.617 ÷ 0.935)	0.0653
53	Hydrazine Derivatives	6	1	0.778 (0.524 ÷ 0.989)	0.1995
54	Carboxylic Acid Amides	9	2	0.769 (0.55 ÷ 0.964)	0.1624
55	Specific Imine and Thione Derivatives	8	2	0.750 (0.516 ÷ 0.959)	0.2183
56	alpha,omega-Dihaloalkanes	2	0	0.750 (0.368 ÷ 1.000)	0.3947
57	Aminoacridine DNA Intercalators	2	0	0.750 (0.368 ÷ 1.000)	0.3947
58	Bipyridilium Herbicides	2	0	0.750 (0.368 ÷ 1.000)	0.3947
59	Cyanohydrins	2	0	0.750 (0.368 ÷ 1.000)	0.3947

60	Ethenyl Pyridines	2	0	0.750 (0.368 ÷ 1.000)	0.3947
61	Hexahydrotriazine Derivatives	2	0	0.750 (0.368 ÷ 1.000)	0.3947
62	N-Haloacylamides	2	0	0.750 (0.368 ÷ 1.000)	0.3947
63	Propargyl Derivatives	2	0	0.750 (0.368 ÷ 1.000)	0.3947
64	Propyne Derivatives	2	0	0.750 (0.368 ÷ 1.000)	0.3947
65	Dicarbonyl Compounds	7	2	0.727 (0.478 ÷ 0.954)	0.2895
66	Sterically Hindered Piperidine Derivatives	4	1	0.714 (0.409 ÷ 0.982)	0.3876
67	Isocyanates and Diisocyanates	6	2	0.700 (0.432 ÷ 0.946)	0.3775
68	Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids	10	4	0.688 (0.467 ÷ 0.895)	0.3580
69	Acridone, Thioxanthone, Xanthone and Phenazine Derivatives	3	1	0.667 (0.330 ÷ 0.974)	0.5243
70	Nitroarenes with Other Active Groups	3	1	0.667 (0.330 ÷ 0.974)	0.5243
71	Bleomycin and Structurally Related Chemicals	1	0	0.667 (0.224 ÷ 1.000)	0.6281
72	DNA Intercalators with Carboxamide and Aminoalkylamine Side Chain	1	0	0.667 (0.224 ÷ 1.000)	0.6281
73	Flavonoids	1	0	0.667 (0.224 ÷ 1.000)	0.6281
74	Four-and Five-Membered Lactones	1	0	0.667 (0.224 ÷ 1.000)	0.6281
75	Halofuranones	1	0	0.667 (0.224 ÷ 1.000)	0.6281
76	N-Acetoxyamines	1	0	0.667 (0.224 ÷ 1.000)	0.6281
77	N-Aryl-N-Acetoxy(Benzoyloxy) Acetamides	1	0	0.667 (0.224 ÷ 1.000)	0.6281
78	Nitroalkanes	1	0	0.667 (0.224 ÷ 1.000)	0.6281

79	Non-Cyclic Alkyl Phosphoramides and Thionophosphoramides	1	0	0.667 (0.224 ÷ 1.000)	0.6281
80	N-Oxicarbonyl amides, N-Acyloxy-N-alkoxyamides	1	0	0.667 (0.224 ÷ 1.000)	0.6281
81	Organic Azides	1	0	0.667 (0.224 ÷ 1.000)	0.6281
82	Pyrrolizidine derivatives	1	0	0.667 (0.224 ÷ 1.000)	0.6281
83	Quinoline Derivatives	1	0	0.667 (0.224 ÷ 1.000)	0.6281
84	Specific 5-Substituted Uracil Derivatives	1	0	0.667 (0.224 ÷ 1.000)	0.6281
85	Specific Acetate Esters	1	0	0.667 (0.224 ÷ 1.000)	0.6281
86	Sultones	1	0	0.667 (0.224 ÷ 1.000)	0.6281
87	Sultones protein binding	1	0	0.667 (0.224 ÷ 1.000)	0.6281
88	Thiols	1	0	0.667 (0.224 ÷ 1.000)	0.6281
89	alfa,beta-Unsaturated aldehydes	4	2	0.625 (0.315 ÷ 0.919)	0.6022
90	Hydroxamic acid	4	2	0.625 (0.315 ÷ 0.919)	0.6022
91	Nitroaniline Derivatives	12	7	0.619 (0.416 ÷ 0.816)	0.5882
92	(Thio)Acyl and (thio)carbamoyl halides, cyanides, azides, etc.	11	7	0.600 (0.391 ÷ 0.804)	0.6575
93	Amino Anthraquinones	2	1	0.600 (0.228 ÷ 0.956)	0.6877
94	Nitro Azoarenes and p-Monosubstituted Azobenzene Derivatives	2	1	0.600 (0.228 ÷ 0.956)	0.6877
95	Geminal Polyhaloalkane Derivatives	10	7	0.579 (0.363 ÷ 0.790)	0.7250
96	Haloalkene Derivatives with Electron-Withdrawing Groups	4	3	0.556 (0.254 ÷ 0.851)	0.7617
97	Diazenes and Azoxyalkanes	1	1	0.500 (0.094 ÷ 0.906)	0.8614

98	N,N-Dialkyldithiocarbamate Derivatives	1	1	0.500 (0.094 ÷ 0.906)	0.8614
99	p-Substituted Mononitrobenzenes	1	1	0.500 (0.094 ÷ 0.906)	0.8614
100	Monohaloalkanes	1	3	0.333 (0.026 ÷ 0.670)	0.9806
101	Polynitroarenes	1	3	0.333 (0.026 ÷ 0.670)	0.9806
102	Arenediazonium Salts	0	1	0.333 (0.000 ÷ 0.776)	1.0000
103	Organic Peroxy Compounds	0	1	0.333 (0.000 ÷ 0.776)	1.0000
104	Quinolone Derivatives	0	1	0.333 (0.000 ÷ 0.776)	1.0000
105	p-Aminobiphenyl Analogs	0	2	0.250 (0.000 ÷ 0.632)	1.0000

Alerts which are not supported by chemicals from the training set (theoretical alerts) are not included in Table 1. Detailed information for each alert such as structural boundaries, mechanisms, local training sets and references associated with each observed data is provided in ***Appendix 2***.

4.5. Algorithm and descriptor generation

The structural boundaries of the alerts are derived from the chemicals included in the local training sets (see Section 4.3). For derivation of each alert mechanistically justifiable structural fragments for interaction with DNA and/or proteins 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

TIMES Chromosomal aberrations model version 17.17

4.7. Chemicals/Descriptors ratio

Provided in Section 4.4.

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.1.12 (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

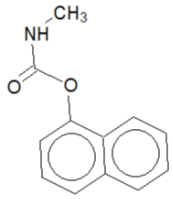
Applicability domain of the CA model (+S9) include three sub-domain layers: general parametric requirements, structural features and alerts reliability.

- 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: 63-25-2
- Name: Carbaryl
- 2D Depiction:



Property	Domain	Example chemical

log K_{OW}	[-12.3850; 29.6790]	2.348
MW, Da	[44.006; 1514.0861]	201.211

* K_{OW} is calculated by EPI Suite

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 extracted from 855 training chemicals contains:

  - 2375 correct fragments,
  - 220 fuzzy fragments (treated as correct fragments),
  - 246 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).

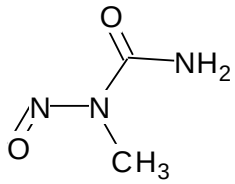
Example chemical belonging to alert with “High reliability”.

Chemical ID:

CAS: 684-93-5

Name: 1-methyl-1-nitrosourea

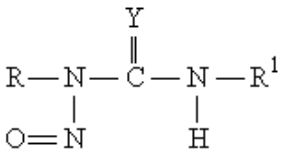
2D depiction:



Belonging to alert:

Name: Alkylated nitrosoureas and nitrosoguanidines

Structural boundaries:



where:

R = (Csp^3)n – acy at n ≥ 1, preferably linear or branched C1–C5 alkyl groups; R may also include an acyl group C(=O)Csp^3 (acy) and a nitro group;

R1 = H atom; (Csp^3)n – acy at n ≥ 1; C(=O)Csp^3 (acy), Csp^2 (aryl), nitro group, etc.;

Y = O and NH.

Reliability:

"High reliability" based on AP=1; N=18 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_CA model (+S9) includes 855 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

Descriptors in the models are structural alerts. The main characteristics of each alert are provided in Table 1 (Section 4.4).

6.4. Data for the dependent variable for the training set

The training set of 855 chemicals include:

- 538 chemicals having positive observed CA data
- 317 chemicals having negative observed CA data.

Distribution of positive/negative chemicals in the training set of model is used for estimating performance and confidence range of the so-called *naïve alert* which is 0.628 (0.596 ÷ 0.660).

6.5. Other information about the training set

The training set is compiled according to the recommendations described in the OECD TG473.

6.6. Pre-processing of data before modelling

Not available.

6.7. Statistics for goodness-of-fit

During the internal validation the original training set is separated many times randomly into two parts – one becomes a training set and the other becomes a test set. The model is re-derived many times using each new training set. Then, performance is estimated for the training sets and test sets. The averaged value of all training set performances is compared to the averaged value of all test set performances in order to assess the amount of optimism in the goodness-of-fit (GOF optimism) in the original model. GOF optimism is calculated as average performance over training sets minus average performance over test sets. Results are provided in Table 2.

Table 2. Performance of the original model over its training set (goodness-of-fit, GOF) vs. expected performance over set different from the training set (GOF – GOF optimism).

	Performance $_{est.}$[1)] model	*p-value*[1)]	Performance[2) 3)] different set
All predictions (accuracy)	0.820 (0.794 ÷ 0.845)	$< 10^{-10}$	0.808
Positive chemicals (sensitivity)	0.861 (0.831 ÷ 0.889)	$< 10^{-10}$	0.838
Negative chemicals (specificity)	0.749 (0.701 ÷ 0.796)	$< 10^{-10}$	0.748

[1]) Confidence ranges and *p-value* are calculated at 95% confidence level

[2]) Estimated performance for training set minus GOF optimism calculated from internal validation

[3]) Estimation of expected performance over external sets (different from training sets)

Addition information including mathematical formalism underlying the above statistical measures are provided in ***Appendix 3***.

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

Not performed

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

Method 1. *k-fold* cross-validation

In ***k-fold cross-validation*** the original training set is partitioned into *k* equally sized subsets. Each time a single subset is used as a test set and the remaining *k-1* subsets are used as training set. In this manner the process is repeated *k* times and each data from the original training set is used once as a test data and *k-1* times as a training data. The advantage of this method is that any data is used for both training and validation and each data is used exactly once as a test data. Commonly the 10-fold cross-validation is used (90% training data, 10% test data). In addition, 4-fold cross validation (75% training data, 25% test data) is also performed and the results from both procedures are provided in Table 3.

Table 3. Results from *k-fold* (*10-fold* and *4-fold*) cross-validation.

	10-fold	**4-fold**

	Training sets	Test sets	Training sets	Test sets
Unique chemicals, %	90.0 (89.9 ÷ 90.1)	10.0 (9.9 ÷ 10.1)	75.0 (74.9 ÷ 75.1)	25.0 (24.9 ÷ 25.1)
Performance $_{est.}$, all predictions (accuracy)	0.820 (0.810 ÷ 0.830)	0.806 (0.720 ÷ 0.892)	0.820 (0.814 ÷ 0.825)	0.807 (0.774 ÷ 0.840)
p-value, accuracy	$< 10^{-10}$	0.0002	$< 10^{-10}$	$< 10^{-10}$
Performance $_{est.}$, positive chemicals (sensitivity)	0.860 (0.850 ÷ 0.871)	0.837 (0.740 ÷ 0.934)	0.860 (0.838 ÷ 0.882)	0.836 (0.744 ÷ 0.929)
p-value, sensitivity	$< 10^{-10}$	0.0226	$< 10^{-10}$	0.0001
Performance $_{est.}$, negative chemicals (specificity)	0.749 (0.726 ÷ 0.772)	0.736 (0.537 ÷ 0.935)	0.748 (0.702 ÷ 0.794)	0.743 (0.612 ÷ 0.874)
p-value, specificity	$< 10^{-10}$	0.0151	$< 10^{-10}$	8.7×10^{-8}

[1]) Confidence ranges and *p-value* are calculated at 95% confidence level

Method 2. *Monte Carlo* cross-validation

In ***Monte Carlo cross-validation*** the original training set is split randomly into training and test set. The advantage of this method (compared to *k-fold* cross validation) is that the proportion between training and test sets does not depend on the number of repetitions in the internal validation procedure. The ***Monte Carlo cross-validation*** (similarly to the *bootstrapping*) suppose creating a large number of new training/test sets (1000 – 10000). Results from application of this statistical method are provided in Table 5.

Table 4. Results from Monte Carlo cross-validation (1000 repetitions).

	75% training set		63% training set	
	Training sets	Test sets	Training sets	Test sets
Unique chemicals, %	75.0 (75.0 ÷ 75.0)	25.0 (25.0 ÷ 25.0)	63.0 (63.0 ÷ 63.0)	37.0 (37.0 ÷ 37.0)
Performance $_{est.}$, all predictions (accuracy)	0.820 (0.804 ÷ 0.835)	0.808 (0.761 ÷ 0.856)	0.819 (0.799 ÷ 0.839)	0.808 (0.773 ÷ 0.843)
p-value, accuracy	$< 10^{-10}$	1.7×10^{-9}	$< 10^{-10}$	$< 10^{-10}$
Performance $_{est.}$, positive chemicals (sensitivity)	0.860 (0.843 ÷ 0.878)	0.839 (0.784 ÷ 0.894)	0.860 (0.838 ÷ 0.882)	0.839 (0.799 ÷ 0.878)
p-value, sensitivity	$< 10^{-10}$	0.0003	$< 10^{-10}$	7.1×10^{-6}

Performance $_{est.}$, negative chemicals (specificity)	0.748 (0.720 ÷ 0.777)	0.749 (0.664 ÷ 0.834)	0.748 (0.710 ÷ 0.786)	0.751 (0.687 ÷ 0.815)
p-value, specificity	$< 10^{-10}$	9.3×10^{-7}	$< 10^{-10}$	7.6×10^{-10}

[1]) Confidence ranges and *p-value* are calculated at 95% confidence level

6.10. Robustness - Statistics obtained by Y-scrambling

Not performed

6.11. Robustness - Statistics obtained by bootstrap

In bootstrapping a newly derived training sets is populated from the original training set of the model by random sampling with replacement until the size of the new training set reaches the size of the original training set. The data not selected for the new training set becomes the new test set. On average, about 63% of original training set data goes into the new training set (some data appear more than once) and 37% remains in the new test set. One of the advantages of this method is that the new training sets and the original training set are equally sized. The process is repeated many times and the average results are provided in Table 5.

Table 5. Results from bootstrapping (1000 repetitions).

	Training sets	**Test sets**
Unique chemicals, %	63.2 (61.1 ÷ 65.3)	36.8 (34.6 ÷ 38.9)
Performance $_{est.}$, all predictions (accuracy)	0.819 (0.793 ÷ 0.845)	0.808 (0.773 ÷ 0.844)
p-value, accuracy	$< 10^{-10}$	$< 10^{-10}$
Performance $_{est.}$, positive chemicals (sensitivity)	0.861 (0.831 ÷ 0.891)	0.838 (0.796 ÷ 0.881)
p-value, sensitivity	$< 10^{-10}$	6.8×10^{-6}
Performance $_{est.}$, negative chemicals (specificity)	0.747 (0.699 ÷ 0.795)	0.753 (0.691 ÷ 0.815)
p-value, specificity	$< 10^{-10}$	6.3×10^{-10}

[1]) Confidence ranges and *p-value* are calculated at 95% confidence level

6.12. Robustness - Statistics obtained by other methods

Not performed

6.13. Comment on the internal validation of the model

The model shows good predictive performance for positive chemicals (Sensitivity) – 86% for training set, around 84% expected Sensitivity for chemicals different from the training set. On the contrary, prediction performance for negative chemicals (Specificity) is not that high (~75%) indicating for possible over-prediction. Lower Specificity could be due to the fact that the model training set is relatively small and not balanced – 60% of the training chemicals are observed positive which favours training of positive chemicals rather than negative ones. One should also address that *in vitro* CA tests is known to provide false positive predictions due to cytotoxicity.

Section 7. External validation – OECD Principle 4

7.1. Availability of the external validation set

93 external chemicals are available to examine performance of the model.

7.2. Available information for the external validation set

The external validation set includes pesticide chemicals from the EFSA database. Details of the EFSA database are available in the website of the agency: https://data.europa.eu/euodp/en/data/dataset/database-pesticide-genotoxicity-endpoints

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 EFSA genotoxicity database is biased towards negative chemicals. This justifies the fact that Sensitivity of the model is examined based on a small number of chemicals as

compared with the substances used for estimating Specificity of the model. Details about the external validation set are available in:

P.I. Petkov, T. W. Schultz, M. Honma, T. Yamada, E. Kaloyanova, O. Mekenyan. 2019. Validation of the performance of TIMES genotoxicity models with EFSA pesticide data. *Mutagenesis*, Vol. 34, pp. 83-90.

7.6. Experimental design of test set

The external validation set of 93 pesticides contains:

- Positive data (10 chemicals)
- Negative data (83 chemicals)

7.7. Predictivity – Statistics obtained by external validation

Performance of the chemicals which belong and does not belong to the model domain (*In domain* and *Out of domain*) is:

- Sensitivity = 70% (7 of 10 chemicals)
- Specificity = 43% (36 of 83 chemicals)

Lower Specificity indicated that the *in vitro* liver S9 metabolic simulator associated with the model has to be modified with respect to the pesticide chemicals. These and other modifications of the metabolic simulator are performed in the latest version of the model.

7.8. Predictivity – Assessment of the external validation set

The EFSA genotoxicity database for pesticide chemicals is assumed to be a high quality database. According to the description of the database, all endpoints are evaluated using standard tests conditions as described in the corresponding tests guidelines.

7.9. Comment on the external validation of the model

Performance of the CA model (+S9) in terms of Sensitivity is reasonable accounting for the fact that all 10 chemicals with positive CA data do not belong to the model domain. Lower Specificity indicates that modifications of some alerts and/or *in vitro* S9 metabolic simulator associated with the model are needed. Such modifications are performed in the latest versions of the model.

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. Fitting of model variable to the observed data,
d. Verification of model quality,
e. Depending on the results found in step *d* model building could continue with step *a*, *b* or *f*,
f. 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 analyze 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.