	QMRF identifier: TOXFENCE-QMRF-004a
	QMRF Title: Clastogenicity in vivo Micronucleus Test TOXFENCE model – version 1.0
	Date of QMRF: 25/03/2026
	Model Developer: RAMC Co., Ltd.

1. QSAR identifier

1.1 QSAR identifier (title)

Clastogenicity in vivo Micronucleus Test TOXFENCE model – version 1.0

1.2 Other related models

Mutagenicity AMES test with S9 TOXFENCE model – version 1.0

Mutagenicity AMES test without S9 TOXFENCE model – version 1.0

Clastogenicity in vitro Chromosomal Aberration test TOXFENCE model – version 1.0

1.3 Software coding the model

TOXFENCE v1.0

The model is implemented in TOXFENCE, a web-based SaaS platform developed by Risk Management Consulting Co., Ltd. TOXFENCE is designed to perform QSAR-based toxicity prediction using chemical structure information as input. The software is provided as an online service without local installation and includes functions for model execution, result review, and report output. The model was implemented in a Python-based environment, the backend service is operated using FastAPI, and RDKit was used for molecular structure handling and descriptor generation.

Risk Management Consulting Co., Ltd.

https://www.toxfence.com

2 General information

2.1 Date of QMRF

Feb 2025

2.2 QMRF author(s) and contact details

[1] Organisation: Risk Assessment & Management Consulting 04156 Seoul, Korea

[2] Contact e-mail: ramc0983@naver.com

[3] Corporate website: https://www.ramc0983.com/

[4] TOXFENCE web service: https://www.toxfence.com/

2.3 Date of QMRF update(s)

25/03/2026

2.4 QMRF update(s)

Updated by: Risk Assessment & Management Consulting Co., Ltd.

Contact: ramc0983@naver.com

Modified field: 1.3 Software coding the model

Reason for update: The website information was revised to replace the service address with the company website address.

2.5 Model developer(s) and contact details

The model was developed by Risk Assessment & Management Consulting Co., Ltd., 04156 Seoul, Republic of Korea.

Contact e-mail: ramc0983@naver.com

Website: https://www.ramc0983.com/

TOXFENCE web service: https://www.toxfence.com/

2.6 Date of model development and/or publication

2025

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

[1] scikit-learn 1.5.2: https://scikit-learn.org/stable/

[2] xgboost 2.1.3: https://xgboost.ai/

[3] lightgbm 4.5.0: https://lightgbm.readthedocs.io/en/stable/

2.8 Availability of information about the model

The model is proprietary. The training dataset and the model are not publicly available. The model is implemented and operated through the TOXFENCE web service.

https://www.toxfence.com/

2.9 Availability of another QMRF for exactly the same model

Another QMRF is not available.

3 Defining the endpoint - OECD Principle 1

3.1 Species

Mice, rats, or other appropriate mammalian species

3.2 Endpoint

in vivo Micronucleus Test (chromosome aberration)

3.3 Comment on endpoint

Genotoxicity is a broad term that encompasses processes which alter the structure, information content, or segregation of DNA, and is not necessarily linked to mutagenicity. Therefore, genotoxicity tests include those that indicate DNA damage (such as DNA strand breaks, unscheduled DNA synthesis (UDS), sister chromatid exchange (SCE), DNA adduct formation, or mitotic recombination) without providing direct evidence of mutation, as well as tests for mutagenicity. The mammalian in vivo micronucleus test is particularly relevant for assessing genotoxicity because it involves factors like in vivo metabolism, pharmacokinetics, and DNA repair processes, which, although species-dependent, actively contribute to the observed responses. Furthermore, in vivo assays are valuable for further investigating genotoxicity signals detected in in vitro systems [2].

3.4 Endpoint units

Dimensionless

3.5 Dependent variable

The dependent variable is classified as genotoxic or non-genotoxic.

3.6 Experimental protocol

The mammalian in vivo micronucleus test is employed to detect damage caused by the test substance to the chromosomes or mitotic apparatus of erythroblasts. This test assesses the formation of micronuclei in erythrocytes, which are collected from either bone marrow or peripheral blood cells of animals, typically rodents.

3.7 Endpoint data quality and variability

Experimental data for this model were curated from studies accessible through eChemPortal and the Hazardous Substances Data Bank (HSDB). Only experimental data with reliability scores of 1 or 2 were retained for model development. The dataset was restricted to organic compounds tested in the in vivo mammalian erythrocyte micronucleus test, based on results obtained from erythrocytes sampled from bone marrow and/or peripheral blood of test animals. During data curation, salts and inorganic substances were removed, duplicate records were excluded, and conflicting results reported for the same substance were resolved through expert consultation. For each retained substance, CAS number, SMILES, and a final binary outcome classified as positive (1) or negative (0) were recorded. SMILES were standardised using a custom Python 3.10.4 script based on the RDKit library. After curation, the final dataset consisted of 2,386 organic compounds. As the data were compiled from multiple studies and sources, some inter-study and inter-laboratory variability may remain due to differences in test conditions, animal species, sampling tissues, dosing regimens, and study reporting.

4 Defining the algorithm - OECD Principle 2

4.1 Type of model

A model in which multiple tree-based machine learning algorithms soft-vote the predicted result to select the final predicted value.

4.2 Explicit algorithm

The Voting model was constructed with Random Forest, XGBoost, Light GBM, and Histogram Gradient Boosting models. The hyperparameter value of each model was determined by Grid Search. Each model determines whether or not it is toxic as a value between 0 and 1, and classifies it by arithmetic average.

4.3 Descriptors in the model

1) Chi1n

The first-order molecular connectivity index based on the number of non-hydrogen atoms, representing simple branching characteristics of the molecule.

2) HallKierAlpha

Adescriptor related to the alpha parameter in the Hall–Kier molecular connectivity theory, reflecting molecular branching and flexibility.

3) FractionCSP3

The fraction of carbon atoms in sp^3 hybridization state within the molecule, indicating the degree of molecular saturation.

4) Chi3n

The third-order molecular connectivity index based on non-hydrogen atoms, capturing more complex branching and molecular size characteristics.

5) VSA_EState1

The sum of Electrotopological State (EState) values for atoms in a molecule within a specific van der Waals surface area range (Bin 1), reflecting electronic and steric properties.

6) Chi3v

The third-order valence molecular connectivity index, considering both atom types and bonding information to describe molecular complexity.

7) PEOE_VSA11

The sum of atomic partial charges calculated by Partial Equalization of Orbital Electronegativities (PEOE) for atoms in a specific van der Waals surface area range (Bin 11), reflecting charge distribution and molecular polarity.

8) VSA_EState9

The sum of Electrotopological State (EState) values for atoms in a molecule within van der Waals surface area range Bin 9, describing specific size and electronic environment contributions.

9) NumValenceElectrons

The total number of valence electrons in the molecule, representing the molecule's potential for forming bonds and undergoing reactions.

10) NumRotatableBonds

The number of rotatable single bonds in the molecule, reflecting molecular flexibility and conformational variability.

11) Chi4n

The fourth-order molecular connectivity index based on non-hydrogen atoms, describing even more complex branching and topological features.

12) FpDensityMorgan2

The density of Morgan fingerprint features generated with radius 2, representing the richness of local atomic environments in the molecule.

13) MinAbsPartialCharge

The minimum absolute value of atomic partial charges in the molecule, reflecting the least polarized site with respect to charge separation.

14) EState_VSA

The volume-based surface area calculated using the Electrotopological State (EState) index, reflecting the distribution of electronic properties across the molecule.

15) TPSA

The topological polar surface area, calculated based on polar fragments, representing the molecule's ability to interact through hydrogen bonding.

16) PEOE_VSA4

The sum of atomic partial charges (calculated using PEOE) for atoms in van der Waals surface area Bin 4, providing information about medium-polarity atom regions.

17) fr_unbrch_alkane

The number of unbranched alkane fragments (linear chains of carbon atoms) present in the molecule, reflecting simple hydrocarbon structure.

4.4 Descriptor selection

Molecular descriptors were calculated using the RDKit cheminformatics library within a custom Python 3.10.4 workflow. A total of 208 molecular descriptors were calculated, encompassing various physicochemical and structural properties of the compounds. From this comprehensive set, descriptor selection was based on SHAP (SHapley Additive exPlanations) values, prioritizing the removal of descriptors with low importance. The optimal set of 18 descriptors was finalized by comparing model performance before and after descriptor removal, selecting those that improved model performance. The selection process employed a subset evaluator that utilized Stratified 5-fold cross-validation scores on the training data set.

4.5 Algorithm and descriptor generation

The RDKit Descriptor Calculation module in Python 3.10.4 was utilized to compute a comprehensive set of molecular descriptors, encompassing various physicochemical and structural properties of the compounds.

4.6 Software name and version for descriptor generation

Molecular descriptors were generated using RDKit 2023.09 in Python 3.10.4.

https://www.rdkit.org/

https://www.python.org/

4.7 Chemicals/Descriptors ratio

1,851 chemicals / 18 descriptors = 102.8

5 Defining the applicability domain - OECD Principle 3

5.1 Description of the applicability domain of the model

The applicability domain is defined in the descriptor space used by the model. AD is assessed by combining Mahalanobis distance squared (MD2) and the k-nearest-neighbor (k-NN) mean distance in the standardized feature space. A query is considered inside the AD only when both criteria are met.

5.2 Method used to assess the applicability domain

Training descriptors are scaled using StandardScaler with optional z-clipping where applicable. In the standardized space, MD2 is computed using a LedoitWolf shrinkage covariance model and the 95th percentile of the training MD2 distribution is used as the threshold. In the same space, the k-NN mean distance is computed and the 95th percentile of the training k-NN mean-distance distribution is used as the threshold. A query is classified as inside the AD only when both thresholds are satisfied. Otherwise, it is classified as outside the AD.

5.3 Software name and version for applicability domain assessment

Implemented in TOXFENCE v1.0.

Applicability domain assessment was performed in the standardized descriptor space using Mahalanobis distance squared (MD2) based on a LedoitWolf shrinkage covariance model and k-nearest neighbor mean distance. A query was classified as inside the applicability domain only when both the MD2 threshold and the k-NN mean-distance threshold, defined as the 95th percentile of the corresponding training distributions, were satisfied.

5.4 Limits of applicability

This model is intended for organic chemicals that can be represented by valid SMILES. For salts and multi-component substances, descriptors are calculated using a parent structure obtained by removing inorganic counterions and selecting the main organic fragment. Therefore, the prediction reflects the processed parent structure rather than the full mixture or salt form. Predictions for purely inorganic substances or chemotypes that are underrepresented in the training data may have increased uncertainty and should be interpreted conservatively together with the applicability domain result.

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:No

INChI: No

MOL file: Yes

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

The training set consisted of 1,851 organic substances, including 713 positive and 1,138 negative substances. The compounds included in the training set were selected after data curation according to the predefined inclusion criteria for the endpoint. The final training set was constructed to include curated organic compounds with reliable binary classifications suitable for model development.

6.6 Pre-processing of data before modelling

To reduce the effect of class imbalance, the training procedure incorporated class weighting so that underrepresented positive substances were given greater importance during model development. No resampling was performed, and the original number of training substances was retained.

6.7 Statistics for goodness-of-fit

Accuracy = 84%, Precision = 65%

Recall = 89%, Specificity = 81%

F1-score = 75%, MCC = 66%

TP 372, TN 872, FP 194, FN 42

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

Accuracy = 68 %, Precision = 45%

Recall = 64%, Specificity = 70%

F1-score = 53%, MCC = 31%

TP: 67, TN: 187, FP: 80, FN: 37

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: No

INChI: No

MOL file: Yes

7.3 Data for each descriptor variable for the external validation set

All

7.4 Data for the dependent variable for the external validation set

All

7.5 Other information about the external validation set

The test set is composed of 463 substances (186 positive, 277 negative)

7.6 Experimental design of test set

Prior to model development, the full dataset was randomly divided into a training set and a test set at a ratio of 8:2.

7.7 Predictivity - Statistics obtained by external validation

Accuracy = 72%, Precision = 50%

Recall = 72%, Specificity = 72%

F1-score = 59%, MCC = 41%

TP 94, TN 242, FP 92, FN 35

7.8 Predictivity - Assessment of the external validation set

The external validation set was considered appropriate for assessing model predictivity because it was sufficiently large and was generated by random splitting prior to model development. The test set was assumed to be broadly representative of the descriptor space and response distribution of the training set. Therefore, the external validation results were considered suitable for evaluating the predictive performance of the model within its applicability domain.

7.9 Comments on the external validation of the model

The external validation was performed using a hold-out test set obtained by random splitting of the full dataset prior to model development. Therefore, the validation can be regarded as an independent test of the final model, although the test compounds originated from the same overall dataset rather than from a completely separate external source. The reported performance should thus be interpreted as evidence of predictive ability within the chemical space represented by the dataset, while predictions for compounds outside the applicability domain or for underrepresented chemotypes should be interpreted with caution.

8 Providing a mechanistic interpretation - OECD Principle 5

8.1 Mechanistic basis of the model

The model is an ensemble classification model based on molecular descriptors encoding structural and physicochemical properties of chemicals relevant to genotoxicity. Its mechanistic basis is indirect, in that chemicals with similar descriptor patterns are assumed to have similar genotoxic potential. Because the final prediction is generated by combining multiple tree-based machine-learning algorithms, a direct one-to-one mechanistic interpretation is limited.

8.2 Other information about the mechanistic interpretation

The mechanistic interpretation of the model was established a posteriori, based on the interpretation of the final set of descriptors and model outputs after model development.

8.3 Other information about the mechanistic interpretation

No additional information is available regarding the mechanistic interpretation.

9 Miscellaneous information

9.1 Comments

No additional comments.

9.2 Bibliography

[1] Drug Discov Today. 2018 Aug;23(8):1538-1546. doi: 10.1016/j.drudis.2018.05.010. Epub 2018 May 8.

[2] OECD Test No. 474, Mammalian Erythrocyte Micronucleus Test

9.3 Supporting information

Not available.