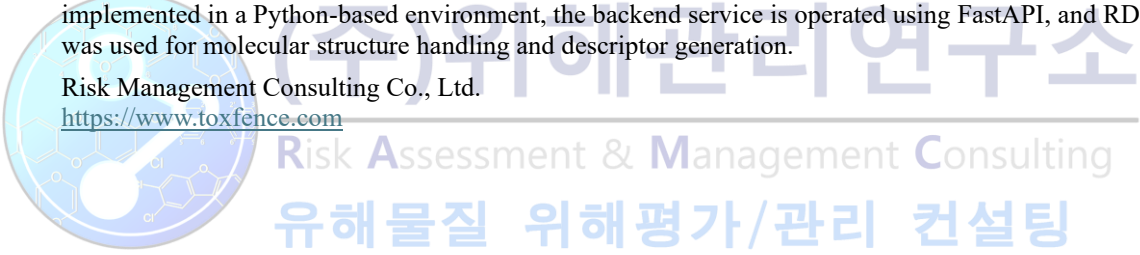




<i>QMRF identifier: TOXFENCE-QMRF-006a</i>
<i>QMRF Title: GHS Skin Corrosion/Irritation TOXFENCE model – version 1.0</i>
<i>Date of QMRF: 25/03/2026</i>
<i>Model Developer: RAMC Co., Ltd.</i>

1. QSAR identifier

- 1.1 QSAR identifier (title)
GHS Skin Corrosion/Irritation TOXFENCE model – version 1.0
- 1.2 Other related models
No related models identified
- 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>



Risk Management Consulting Co., Ltd. Address: #1101, #1401, 331 Dongmak-ro, Mapo-gu, Seoul, Republic of Korea

Email address. ramc0983@naver.com | Tel. 02-715-0983

2 General information

2.1 Date of QMRF

Sep 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.



(주)위해관리연구소

Risk Assessment & Management Consulting

유해물질 위해평가/관리 컨설팅

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) Python 3.11.8

2) scikit-learn 1.8.0

3) xgboost 3.1.1

4) pandas 2.2.3

5) numpy 1.26.4

6) joblib 1.5.3

7) RDkit 2023.09

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

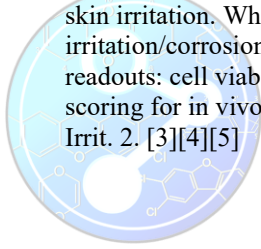
The albino rabbit is the primary in vivo reference species for skin corrosion/irritation testing under OECD TG 404. However, the present model was developed using a dataset constructed according to GHS classification criteria and compiled from available experimental data, rather than from a single species-specific test system. Accordingly, in vitro methods relevant to skin corrosion/irritation classification, including reconstructed human epidermis and related assays (e.g., OECD TG 430, TG 431, TG 435, and TG 439), may also contribute to the underlying evidence where available.

3.2 Endpoint

Skin Corrosion/Irritation

3.3 Comment on endpoint

Skin corrosion/irritation is defined under UN GHS/CLP as the induction of irreversible tissue damage (corrosion) or reversible inflammation (irritation) following dermal exposure. Corrosion is characterized by visible necrosis through the epidermis into the dermis; irritation denotes reversible damage. These hazards are explicitly classified as Skin Corr. 1 (1A/1B/1C) and Skin Irrit. 2 under the harmonized criteria. [1][2] A tiered evaluation strategy is generally recommended: consideration of existing information, followed by validated in vitro methods (preferentially), and resorting to in vivo testing only when necessary. For in vitro, reconstructed human epidermis (RhE) methods are internationally accepted: OECD TG 431 for skin corrosion (with partial sub-categorization) and OECD TG 439 for skin irritation. Where in vivo data are required, OECD TG 404 provides the guideline for acute dermal irritation/corrosion. [3][4][5] These methods align with GHS/CLP decision logic and provide robust readouts: cell viability thresholds for RhE in vitro assays (TG 431/439) and erythema/eschar and edema scoring for in vivo studies (TG 404), enabling consistent GHS classification for Skin Corr. 1 and Skin Irrit. 2. [3][4][5]



Risk Assessment & Management Consulting
유해물질 위해평가/관리 컨설팅

3.4 Endpoint units

Dimensionless

3.5 Dependent variable

The dependent variable was defined as positive (1) for substances classified as skin corrosion or skin irritation according to the GHS classification criteria, and negative (0) for substances not classified for either endpoint.

3.6 Experimental protocol

The underlying experimental data were compiled from studies relevant to GHS skin corrosion/irritation classification. These data may include validated in vitro reconstructed human epidermis (RhE) methods, such as OECD TG 431 for skin corrosion and OECD TG 439 for skin irritation, as well as in vivo dermal studies conducted according to OECD TG 404 where available. In vitro evaluations are generally based on predefined cell-viability thresholds, whereas in vivo evaluations rely on erythema, eschar, and edema scoring. Additional supportive information may also be obtained from methods such as the transcutaneous electrical resistance test (OECD TG 430) and membrane barrier test (OECD TG 435).

3.7 Endpoint data quality and variability

Experimental data for this study were sourced from NITE (Japan), HCIS (Australia), and eChemPortal, which provide curated information relevant to GHS skin corrosion/irritation classification. Where available, source records referred to guideline-compliant studies and methods, including OECD TG 431 for in vitro skin corrosion, OECD TG 439 for in vitro skin irritation, and OECD TG 404 for in vivo dermal corrosion/irritation. Because the dataset was compiled from multiple sources, some variability related to differences in original study conditions, reporting formats, and data availability may remain. During data curation, duplicate records were removed, and when both positive and negative classifications were identified for the same substance, the final label was determined through expert review and advisory discussion. After removal of duplicates, salts, and inorganic substances, a dataset of 4,896 substances was compiled. For each substance, the CAS number, SMILES, and a GHS skin corrosion/irritation-based label were recorded, encoded as 1 for substances classified as Skin Corr. 1 (including sub-categories 1A, 1B, and 1C) and/or Skin Irrit. 2, and 0 for substances classified as Not Classified. SMILES were standardized using a custom Python 3.10.4 script based on the RDKit library, and the data were randomly divided into a training set (80%) and a test set (20%).

4 Defining the algorithm - OECD Principle 2

4.1 Type of model

This model is an XGBClassifier model that uses RDKit-derived molecular descriptors as inputs to predict whether a chemical is GHS skin irritation/corrosion positive or negative.

4.2 Explicit algorithm

The model is an XGBoost-based classification algorithm trained on RDKit-generated molecular descriptors. Chemical structures were represented as canonical SMILES, from which RDKit was used to derive the descriptor set serving as the model input. Prior to fitting, features were standardized using StandardScaler, and an Isolation Forest-based filtering procedure was applied to reduce the impact of outliers in the training data. Model hyperparameters were selected using a cross-validation-based optimization approach. The final model provides predictions for GHS skin irritation/corrosion.

4.3 Descriptors in the model

1) MaxAbsEStateIndex

The maximum absolute electrotopological state (EState) index in the molecule, representing the atom with the strongest combined electronic and topological contribution.

2) HallKierAlpha

A topological descriptor related to molecular shape and branching, based on the Kier-Hall alpha-modified connectivity framework.

3) qed

Quantitative Estimate of Drug-likeness (QED), a composite descriptor summarizing multiple physicochemical properties into a single score.

4) BertzCT

The Bertz complexity index, a graph-based descriptor that quantifies molecular structural complexity.

5) MinAbsEStateIndex

The minimum absolute electrotopological state (EState) index in the molecule, representing the atom with the smallest combined electronic and topological contribution.

6) VSA_EState7

A descriptor combining van der Waals surface area and EState values for atoms falling within a defined EState bin (bin 7).

7) PEOE_VSA1

The sum of van der Waals surface areas of atoms whose partial charges fall within the first PEOE charge bin.

8) TPSA

Topological Polar Surface Area, defined as the sum of surface contributions of polar atoms, mainly oxygen and nitrogen, and commonly used to characterize molecular polarity.

9) PEOE_VSA6

The sum of van der Waals surface areas of atoms whose partial charges fall within the sixth PEOE charge bin.

10) VSA_EState3

A descriptor combining van der Waals surface area and EState values for atoms falling within a defined EState bin (bin 3).

11) SMR_VSA3

Risk Management Consulting Co., Ltd. Address: #1101, #1401, 331 Dongmak-ro, Mapo-gu, Seoul, Republic of Korea

Email address: ramc0983@naver.com | Tel. 02-715-0983

The sum of van der Waals surface areas of atoms whose fragmental molar refractivity contributions fall within the third SMR bin.



(주)위해관리연구소

Risk Assessment & Management Consulting

유해물질 위해평가/관리 컨설팅

Risk Management Consulting Co., Ltd. Address: #1101, #1401, 331 Dongmak-ro, Mapo-gu, Seoul, Republic of Korea

Email address: ramc0983@naver.com | Tel. 02-715-0983

12) SMR_VSA1

The sum of van der Waals surface areas of atoms whose fragmental molar refractivity contributions fall within the first SMR bin.

13) PEOE_VSA9

The sum of van der Waals surface areas of atoms whose partial charges fall within the ninth PEOE charge bin.

14) PEOE_VSA8

The sum of van der Waals surface areas of atoms whose partial charges fall within the eighth PEOE charge bin.

15) MinAbsPartialCharge

The minimum absolute partial atomic charge in the molecule.

16) MinPartialCharge

The most negative partial atomic charge in the molecule.

17) SMR_VSA6

The sum of van der Waals surface areas of atoms whose fragmental molar refractivity contributions fall within the sixth SMR bin.

18) FpDensityMorgan3

A descriptor defined as the density of Morgan fingerprint bits at radius 3, reflecting the structural feature richness of the molecule relative to its size.

4.4 Descriptor selection

The input descriptors were selected from a pool of candidate descriptors computable with RDKit through SHAP-based feature importance analysis. Based on this analysis, the 18 most influential descriptors were finalized as the model input variables and were applied consistently across model development, training, and prediction. Records with missing descriptor values were excluded from the analysis.

4.5 Algorithm and descriptor generation

The model was implemented in Python 3.11.8 using XGBoost 3.1.1. Chemical structures were prepared as canonical SMILES, and molecular descriptors were generated with RDKit 2025.09.3. The input representation is descriptor-based, consisting of an 18-dimensional feature vector composed of RDKit-derived physicochemical and structural descriptors. Prior to training, records with missing descriptor values were removed, and features were standardized using StandardScaler. To mitigate the influence of atypical observations, an Isolation Forest-based outlier filtering step was applied to the training data. The final classifier is an XGBoost gradient-boosted decision tree model (XGBClassifier).

4.6 Software name and version for descriptor generation

Python 3.10.4. (<https://www.python.org/>)

RDKit 2023.09 (<https://www.rdkit.org/>)

4.7 Chemicals/Descriptors ratio

2,893 chemicals / 18-dimensional descriptor-based input feature vector = 160.7

5 Defining the applicability domain - OECD Principle 3

5.1 Description of the applicability domain of the model

The Applicability Domain (AD) is evaluated using a proprietary algorithm integrated into the TOXFENCE platform. The evaluation derives an overall AD index by calculating multiple assessment parameters based on structural similarity and common substructure information between the input chemical and similar chemicals in the reference data (training/test datasets and database). The resulting AD index is used as a supporting indicator to determine whether the input chemical falls within the model's learned chemical space, i.e., whether the prediction can be interpreted with an adequate level of confidence. The AD outcome is reported as IN/OUT, where IN indicates the input is within the applicability domain with relatively higher confidence, and OUT indicates the input is outside the applicability domain and the prediction should be interpreted with caution.

5.2 Method used to assess the applicability domain

The applicability domain (AD) of this model is assessed by evaluating, from multiple structural perspectives, whether the input chemical lies within the chemical space represented by the model development data (training and test sets). First, both the input structure and the structures in the development data are standardized for consistent comparison by removing salts; when multiple fragments are present, the largest fragment is used as the representative structure. Structural similarity between the input chemical and chemicals in the development data is then computed using molecular fingerprint-based metrics (e.g., ECFP-series fingerprints and RDKit fingerprint similarities). Where appropriate, descriptor-based similarity in a standardized feature space (distance/similarity in a scaled feature space) and common substructure (MCS) based indicators are also incorporated. Because the individual similarity measures have different scales, they are percentile-normalized and integrated to produce an overall combined similarity score for each candidate; if the input chemical is an exact structural match to a chemical in the development data, the combined similarity is set to the maximum. In addition, an AD score is calculated by jointly considering the similarity level of the top-ranked candidates (top-K), the extent to which the input structure is covered by these candidates (structural fragment coverage), and the consistency of the associated values (Value) among the top candidates. Finally, predefined minimum similarity criteria are applied to evaluate whether the input chemical is sufficiently close to the chemical space of the development data, and these steps together generate the indicators used for the final AD assessment.

5.3 Software name and version for applicability domain assessment

Applicability domain assessment was implemented internally in TOXFENCE v1.0, using an embedding-based workflow integrated into the model evaluation pipeline.

5.4 Limits of applicability

The applicability of this model is constrained by the chemical space represented in the model development data and is primarily intended for organic chemicals. If an input is provided in a salt or multi-component form, the assessment is performed after a standardization step in which salts are stripped and the largest fragment is used as the representative structure for similarity and AD calculations. However, for non-organic substances, complex mixtures, or chemotypes that are poorly represented in the development data (e.g., structures containing atypical elements or highly unusual scaffolds), both the prediction and the AD interpretation may be less reliable and should be treated with caution.

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
After exclusion of invalid structures during graph construction, a total of 2,893 substances were available for model development, including 1,340 positive and 1,553 negative substances. These data were subsequently divided into training and test sets using a stratified 80:20 split.
- 6.6 Pre-processing of data before modelling
All chemical structure were verified using the Python script program.
- 6.7 Statistics for goodness-of-fit
Accuracy = 89%, Precision = 86%
Recall = 88%, Specificity = 88%
F1-score = 88%, MCC = 79%
TP 1223, TN 1360, FP 117, FN 193
- 6.8 Robustness - Statistics obtained by leave-one-out cross-validation
- 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 894 substances (421 positive, 473 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 = 79%, Precision = 78%

Recall = 76%, Specificity = 81%

F1-score = 77%, MCC = 57%

TP 321, TN 382, FP 91, FN 100

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.

Risk Management Consulting Co., Ltd. Address: #1101, #1401, 331 Dongmak-ro, Mapo-gu, Seoul, Republic of Korea

Email address: ramc0983@naver.com | Tel. 02-715-0983

8 Providing a mechanistic interpretation - OECD Principle 5

8.1 Mechanistic basis of the model

This model is a descriptor-based QSAR classification model for predicting GHS skin irritation/corrosion. Molecular descriptors are computed from canonical SMILES using RDKit and serve as inputs that summarize physicochemical and structural characteristics relevant to skin hazard potential. The final classifier is an XGBoost decision-tree model that captures non-linear associations between the descriptor set and the binary irritation/corrosion outcome. The model does not explicitly simulate biological mechanisms therefore, mechanistic plausibility is evaluated in a supportive manner using post hoc interpretation approaches, including structural alerts (SA).

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

Mechanistic interpretation is provided as supportive post hoc evidence (e.g., structural alerts, pKa/Kp proxies, and decision-path narratives) and should be considered complementary to, not a replacement for, regulatory assessment and additional supporting evidence.

9 Miscellaneous information

9.1 Comments

No additional comments.

9.2 Bibliography

- [1] UN SCEGHS. Chapter 3.2 Skin Corrosion/Irritation — Definitions. (2018).
- [2] EU CLP (Annex I, 3.2). Classification criteria for skin corrosion (1A/1B/1C) and irritation (2).
- [3] OECD TG 431. In vitro Skin Corrosion (Reconstructed Human Epidermis). (Official text/PDF).
- [4] OECD TG 439. In vitro Skin Irritation (Reconstructed Human Epidermis). (Official text/PDF).
- [5] OECD TG 404. Acute Dermal Irritation/Corrosion. (Official text/PDF).

9.3 Supporting information

Not available.