Re-Thinking Experimentation

QSAR Model Reporting Format (QMRF) for iSafeRat® CLASS (Classification Labelling Assessment for Skin Sensitisation) model v1.4

1 QSAR identifier

1.1 QSAR identifier (title)

iSafeRat® CLASS (Classification Labelling Assessment for Skin Sensitisation) model v1.4

1.2 Other related models

iSafeRat® MechoA Premium and iSafeRat KOW model are related to iSafeRat CLASS model.

1.3 Software coding the model

iSafeRat® Desktop version 4.3.14 and later versions
Additional documentation: KREATiS (2024) – iSafeRat® Desktop v4.3 User Guide

2 General information

2.0 Abstract

This model was developed to predict the skin sensitisation potential as determined in a Local Lymph Node Assay (LLNA) conducted according to the OECD Guidelines 429, 442A and 442B (OECD, 2010a; OECD, 2010b; OECD, 2018).
The model is an expert-based decision tree generating a final outcome result based on the results of four independent sub- models:

- Hapten/prohapten detection (based on MechoA) v1.2 (hereafter referred as Hapt v1.2)
- Autoxidation (prehapten detection) v1.0 (hereafter referred as AutOx v1.0)
- Skin penetration v1.0 (hereafter referred as SkinAbs v1.0)
- Expected LLNA positive LLNA+ v1.1 (hereafter referred as LLNA+ v1.1)

Those sub-models are either based on structural alerts or physicochemical property thresholds.

2.1 Date of QMRF

23 April 2025

2.2 QMRF author(s) and contact details

Marie Darracq, Gaspard Levet, Franklin Bauer, Emel Ay-Albrecht, Zlatomir Todorov, Paul Thomas, Carole Charmeau.

KREATiS SAS
ZAC St. Hubert
23, rue du Creuzat
38080 L'ISLE D'ABEAU
France
Tel: +33 (0)6 46 46 42 33
Email: contact@kreatis.eu
Website: www.kreatis.eu

2.3 Date of QMRF update(s)

Table 1: Dates of QMRF updates

Date	QMRF update identifier
02 April 2025	KTS/QMRF/CLASSv1.4/05
02 April 2025	KTS/QMRF/CLASSv1.4/04
27 March 2025	KTS/QMRF/CLASSv1.4/03
20 February 2025	KTS/QMRF/CLASSv1.4/02
07 March 2024	KTS/QMRF/CLASSv1.4/01

2.4 QMRF update(s)

Table 2: Contents of QMRF updates

QMRF update identifier	Content
KTS/QMRF/CLASSv1.4/05	Update of QMRF sections 1.1, 1.3, 2.0, 3.4, 4.1, 4.2, 6.2, 6.3, 6.4, 6.5, and 6.6 (and consequently section 2.1, 2.3, 2.4, 2.9, 10.1 and 10.2)
KTS/QMRF/CLASSv1.4/04	Correction of a sentence in section 1.1, 1.3, 2.6 and 5.2 (and consequently section 2.1, 2.9, 10.1 and 10.2)
KTS/QMRF/CLASSv1.4/03	Update of QMRF section 3.7 regarding origin of the dataset (and consequently section 2.1, 2.9, 10.1 and 10.2)
KTS/QMRF/CLASSv1.4/02	Update of QMRF sections 1.3, 2.0, 2.1, 2.2, 2.3, 2.4, 2.7, 2.8, 2.9, 3.1, 3.5, 3.6, 3.7, 5.1.b, 6.7, 6.12, 7.7, 9.2, 10.1, 10.2
KTS/QMRF/CLASSv1.4/01	New QMRF according to QAF template. Update of the model decision tree. Refinement of the alerts of the Hapten detection sub-model. Extension of the data set, constitution of an external validation set to validate the model. Refinement of the applicability domain.

2.5 Model developer(s) and contact details

Gaspard Levet, Etienne Bourgart, Franklin Bauer, Emel Ay-Albrecht, Zlatomir Todorov, Marie Darracq, Paul Thomas, Carole Charmeau.

KREATiS SAS
ZAC St. Hubert
23, rue du Creuzat
38080 L'ISLE D'ABEAU
France
Tel: +33 (0)6 28 98 45 51
Email: contact@kreatis.eu
Website: www.kreatis.eu

2.6 Date of model development and/or publication

The results presented in this QMRF refer to the version of the model published internally on 7 March 2024 within iSafeRat® Desktop v4.3.14 and later versions.

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

Bauer FJ, Thomas PC, Fouchard SY, Neunlist SJM. *High-accuracy prediction of mechanisms of action using structural alerts*. Comput Toxicol. 2018;7:36-45.

Butina D. *Unsupervised data base clustering based on daylight's fingerprint and tanimoto similarity: A fast and automated way to cluster small and large data sets*. J Chem Inf Comput. 1999;39(4):747-750.

iSafeRat® Desktop v4.0 for High-Accuracy QSAR predictions by KREATiS SAS. https://isaferat.kreatis.eu/.

Karlberg AT, Bergström MA, Börje A, Luthman K, Nilsson JL. *Allergic contact dermatitis-formation, structural requirements, and reactivity of skin sensitizers*. Chem Res Toxicol. 2008;21(1):53-69.

Kennard RW, Stone LA. *Computer aided design of experiments*. Technometrics. 1969;11(1):137-148.

RDKit: Open-source cheminformatics. Released 2022.09.5. https://www.rdkit.org/.

2.8 Availability of information about the model

The training set and the external test set of the model are proprietary and have not been made publicly available. Statistics of the model were published in a poster for Eurotox 2024 edition (Levet *et al.*, 2024). Furthermore, a limited validation set (provided by OECD) is publicly available on our website (www.kreatis.eu).
Any queries concerning the model or its validity should be addressed to contact@kreatis.eu. Furthermore, KREATiS undertakes to provide supplementary information to sponsors or regulatory authorities upon request to demonstrate compliance of our QSARs with good practice.

2.9 Availability of another QMRF for exactly the same model

Previous QMRFs for the same model are available upon request at contact@kreatis.eu (*i.e.* KTS/QMRF/CLASSv1.4/01, KTS/QMRF/CLASSv1.4/02, KTS/QMRF/CLASSv1.4/03 and KTS/QMRF/CLASSv1.4/04).

3 Defining the endpoint - OECD Principle 1

3.1 Species

Predictions are made for the mice strains (CBA/Ca, CBA/J, CBA/JN and BALB/c) which are used in the local lymph node assay (LLNA) studies (OECD Test Guideline 429, OECD Test Guideline 442A and OECD Test Guideline 442B).

3.2 Endpoint

TOX 7.4.1. Skin sensitisation.

3.3 Comment on endpoint

Skin sensitisation potential can be assessed following *in vivo*, *in vitro* and *in silico* tests. The measured parameter depends on the type of study. In a LLNA test, only the induction phase of the skin sensitisation is evaluated by measuring the stimulating index (SI) corresponding to lymphocyte proliferation in ear lymph nodes. If the SI is above the cut-off limit, depending on the lymphocyte proliferation level (*e.g.* SI≥3), the substance is identified as a skin sensitiser. In this case, the test substance concentration needed to produce a SI that is indicative of a positive response is calculated to assess the potency of the test substance, *i.e.* Effective Concentration 3 (EC3) (OECD Test Guideline 429, OECD Test Guideline 442A and OECD Test Guideline 442B). This value estimates the skin sensitisation potency of a tested substance. iSafeRat® CLASS predicts only sensitising potential of LLNA outcomes.

3.4 Endpoint units

None. The current version iSafeRat® CLASS predicts only the skin sensitisation potential of substances (*i.e.,* skin sensitiser/not skin sensitiser/expected positive in LLNA only).

3.5 Dependent variable

Based on the Global Harmonisation System (GHS), binary LLNA outcomes (*i.e.* sensitiser/non-sensitiser) were used for modelling purposes.

3.6 Experimental protocol

The model is based on Local Lymph Node Assay data conducted following the standard test protocol OECD Test Guideline 429, OECD Test Guideline 442A and OECD Test Guideline 442B.

A LLNA test includes negative/positive control and treatment groups of mice (a minimum of 4 animals). The protocol describes the use of CBA mice in general. The test substance is dissolved/suspended in a carefully selected vehicle (*e.g.* acetone: olive oil). Appropriate concentrations of the substance (without systemic toxicity and/or excessive local skin irritation) are applied on each ear of the mice. The scientific basis for the test is the amount of lymphocyte 3H-methyl thymidine measured in the draining lymph nodes of animals topically exposed

at well-defined test substance concentrations. The stimulation index (SI) is estimated through the ratio of thymidine measured in lymph nodes of treated animals and control animals. The test is positive when the stimulation index is at least equal to 3 (SI ≥ 3) resulting in the EC3 derivation.

When a SI ≥ 3 for a higher concentration and SI < 3 for a lower concentration, the calculation formula is:

[equation 1] EC3= c + [[(3−d)/(b−d)] x (a−c)]

where
a = lowest concentration with a stimulation index ≥ 3,
b = actual stimulation index caused by 'a',
c = highest concentration failing to produce a stimulation index of 3,
d = actual stimulation index caused by 'c'.

When SI > 3 for all concentrations and a linear dose-response relationship is found, the extrapolated calculation is:

[equation 2] EC3 = 2^[log2(c) + ((3 -d)/(b-d)) x (log2(a) - log2(c))] (Ryan *et al.*, 2007)

where
a = dose concentration for next to lowest SI above 3,
b = next to lowest SI above 3,
c = dose concentration for lowest SI above 3,
d = lowest SI above 3.

3.7 Endpoint data quality and variability

The experimental skin sensitisation was evaluated using the laboratory protocol established in the OECD TG 429, 442A and 442B. Only data from *in vivo* LLNA studies for skin sensitisation, validated by KREATiS expert toxicologists and classified using the CLP/UN GHS classification system prior to the modelling steps, were used. Data were sourced from various datasets (ECHA, NiceATM, CosUE, and proprietary confidential data). As a result, the data originate from multiple laboratories and span a period of approximately 10 years (*c.a.* 2010–2020). Data curation was conducted as follow:

Physico-chemical parameters and molecular structural data

- Only organic mono-constituents were included in the dataset
- Octanol-water partition coefficient (log K_{OW}, unitless) values were predicted using appropriate QSPR model, and these values were used as input to run the calculations.
- Molecular weight (MW in g/mol) was calculated from the structure and therefore does not require validation.
- SMILES code (unitless) was generated from the structure and does not require validation.

Skin sensitisation: Local lymph node assay (OECD TG 429)

- Strains
  Only results from female mice were used in the QSAR elaboration. The CBA/Ca or CBA/J mice are the recommended species in the OECD TG 429. Studies in which the mice species was defined as "CBA" only were also included when all the other parameters were valid.
- Number of animals
  The OECD TG 429 requires a minimum of 4 female mice. All the studies following this guideline were included in the dataset. Moreover, all studies which used more than 4 animals were also included.
- Number of groups
  The OECD TG 429 requires a minimum of 3 concentrations groups, a negative control group and a positive control group. All the studies following this guideline were included in the dataset. Moreover, all studies which used more than 3 concentration groups were also included.
- Pre-screen test
  All studies performing pre-screen test with one or two animals for 6 days to assess skin irritation/corrosion and systemic toxicity were included in the dataset (OECD TG 429). If no information on the pre-screen test was available, skin irritation/corrosion and systemic toxicity were assessed through other available studies for the test substance following OECD TG 404, 431 and 439 and OECD TG 401 and 402 respectively. Substances were included in the dataset when reliable information were

found. LLNA studies were disregarded whenever the substance was defined as corrosive/irritant to skin and/or toxic and no information in the pre-screen test was available.

- Vehicle
  Only LLNA performed with non-wholly aqueous vehicles were included in the dataset as stated by the OECD TG 429.
- Doses
  All studies with a strong dose justification for the main study were included in the dataset. Relevant justification of the dose is defined based on the test substance physico-chemical parameters as follow:
  - If the substance is a liquid, the maximum tested dose should be 100%, except if skin irritation/corrosion and/or systemic toxicity were observed in the pre-screen test or if only SI score > 3 were retrieved in another reliable LLNA study conducted with the same substance.
  - If the substance is a solid, the maximum tested dose should be the highest soluble concentration within the most suitable vehicle, except if skin irritation/corrosion and/or systemic toxicity were observed in the pre-screen test or if only SI score > 3 were retrieved in another reliable LLNA study conducted with the same substance. To increase the dataset, substances with no clear justification on solubility of the test substance within the vehicle were also included when no skin irritation/corrosivity and no systemic toxicity were observed.
- Exposure
  Only results from studies with an applied dose of 25 µL on the dorsum of each ear for 3 days, followed by 2 days without treatment were included. Animal sacrifice should be done on the 6th day, 5h after injection of 250 µL sterile phosphate-buffered saline (PBS) containing 20 µCi (7.4×105 Bq) of tritiated (3H)-methyl thymidine (or alternatively 2 µCi (7.4×104 Bq) of 125I-iododeoxyuridine and 10-5M fluorodeoxyuridine) into all mice via the tail vein.
- Skin sensitisation classification
  Validation of skin sensitisation classification was conducted as described in the following decision tree (Figure 1). All studies with a non-linear dose-response were not included in the dataset. Equations 1 and 2 for EC3 calculation, previously described in section 3.6, were used to validate EC3 of each positive LLNA study.

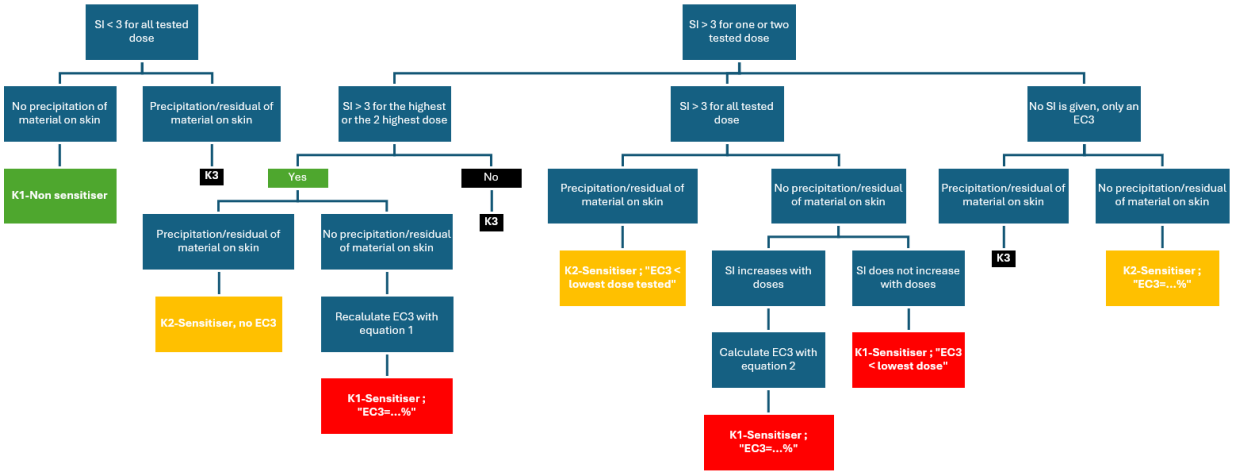

Figure 1: Decision tree to define internal Klimisch scores of the LLNA result when all the other parameters were valid.

- Weight of evidence
  Whenever other *in vivo* and/or *in vitro* studies were available on the same substance (OECD 406, 442C, 442D, 442E), a weight of evidence approach was also considered to strengthen validation of skin sensitisation classification.

Skin sensitisation: Local lymph node assay: DA (OECD TG 442A)

- Strains

Only results from female mice were used in the QSAR elaboration. The CBA/J mice are the recommended species in the OECD TG 442A. Studies in which the mice species was defined as "CBA" only were also included when all the other parameters were valid.

- Number of animals
  The OECD TG 442A requires a minimum of 4 female mice. All the studies following this guideline were included in the dataset. Moreover, all studies which used more than 4 animals were also included.
- Number of groups
  The OECD TG 442A requires a minimum of 3 concentrations groups, a negative control group and a positive control group. All the studies following this guideline were included in the dataset. Moreover, all studies which used more than 3 concentration groups were also included.
- Pre-screen test
  All studies performing pre-screen test with one or two animals for 6 days to assess skin irritation/corrosion and systemic toxicity were included in the dataset (OECD TG 442A). If no information on the pre-screen test was available, skin irritation/corrosion and systemic toxicity were assessed through other available studies for the test substance following OECD TG 404, 431 and 439 and OECD TG 401 and 402 respectively. Substances were included in the dataset when reliable information were found. LLNA studies were disregarded whenever the substance was defined as corrosive/irritant to skin and/or toxic and no information in the pre-screen test was available.
- Vehicle
  Only LLNA performed with non-wholly aqueous vehicle were included in the dataset as stated by the OECD TG 442A.
- Doses
  All studies with a proper dose justification for the main study were included in the dataset. Clear justification of the dose is defined as follows based on the test substance physico-chemical parameters:
    - If the substance is a liquid, the maximum tested dose should be 100%, except if skin irritation/corrosion and/or systemic toxicity were observed in the pre-screen test or if only SI score > 1.8 were retrieved in another reliable LLNA-DA study conducted with the same substance.
    - If the substance is a solid, the maximum tested dose should be the highest soluble concentration within the most suitable vehicle, except if skin irritation/corrosion and/or systemic toxicity were observed in the pre-screen test or if only SI score > 1.8 were retrieved in another reliable LLNA-DA study conducted with the same substance. To increase the dataset, substances with no clear justification on solubility of the test substance within the vehicle were also included when no skin irritation/corrosivity and no systemic toxicity were observed.
- Exposure
  Only results from studies following the experimental schedule of OECD TG 442A were included. Each animal should have been exposed to 1% sodium lauryl sulfate aqueous solution to the dorsum of each ear. One hour after, 25 µL of the test substance should have been applied on the dorsum of each ear for 3 days, followed by 3 days without treatment and a last 7^{th} day of treatment were included. Animal sacrifice should be done on the 8^{th} day.
- Skin sensitisation classification
  Validation of skin sensitisation classification was conducted as described in the following decision tree (Figure 1) with the only difference that bioluminescence ATP content SI score > 1.8, is considered as positive results and no EC1.8 were calculated based on those results. All studies with a non-linear dose-response were not included in the dataset.
- Weight of evidence
  Whenever other *in vivo* and/or *in vitro* studies were available on the same substance (OECD 406, 442C, 442D, 442E), a weight of evidence approach was also considered to strengthen validation of skin sensitisation classification.

Skin sensitisation: Local lymph node assay: BrdU-ELISA or BrdU-FCM (OECD TG 442B)

- Strains
  Only results from female mice were used in the QSAR elaboration. The CBA/JN mice (BrdU-ELISA) or BALB/c mice (BrdU-FCM) are the recommended species in the OECD TG 442B. For BrdU-ELISA, studies

in which the mice species was defined as "CBA" only were also included when all the other parameters were valid.

- Number of animals
  The OECD TG 442B requires a minimum of 4 female mice. All the studies following this guideline were included in the dataset. Moreover, all studies which used more than 4 animals were also included.
- Number of groups
  The OECD TG 442B requires a minimum of 3 concentrations groups, a negative control group and a positive control group. All the studies following this guideline were included in the dataset. Moreover, all studies which used more than 3 concentration groups were also included.
- Pre-screen test
  All studies performing pre-screen test with one or two animals for 6 days to assess skin irritation/corrosion and systemic toxicity were included in the dataset (OECD TG 442B). If no information on the pre-screen test was available, skin irritation/corrosion and systemic toxicity were assessed through other available studies for the test substance following OECD TG 404, 431 and 439 and OECD TG 401 and 402 respectively. Substances were included in the dataset when reliable information were found. LLNA studies were disregarded whenever the substance was defined as corrosive/irritant to skin and/or toxic and no information in the pre-screen test was available.
- Vehicle
  Only LLNA performed with non-wholly aqueous vehicle were included in the dataset as stated by the OECD TG 442B.
- Doses
  All studies with a proper dose justification for the main study were included in the dataset. Clear justification of the dose is defined as follow based on the test substance physico-chemical parameters:
    - If the substance is a liquid, the maximum tested dose should be 100%, except if skin irritation/corrosion and/or systemic toxicity were observed in the pre-screen test or if only SI score > 1.6 (BrdU-ELISA method) or SI score > 2.7 (BrdU-Flow Cytometry Method) were retrieved in another reliable LLNA study conducted with the same substance.
    - If the substance is a solid, the maximum tested dose should be the highest soluble concentration within the most suitable vehicle, except if skin irritation/corrosion and/or systemic toxicity were observed in the pre-screen test or if only SI score > 1.6 (BrdU-ELISA method) or SI score > 2.7 (BrdU-Flow Cytometry Method) were retrieved in another reliable LLNA study conducted with the same substance. To increase the dataset, substances with no clear justification on solubility of the test substance within the vehicle were also included when no skin irritation/corrosivity and no systemic toxicity were observed.
- Exposure
  Only results from studies with an applied dose of 25 µL on the dorsum of each ear for 3 days, followed by one day without treatment were included. Animal sacrifice should be done on the 6^{th} day, one day after intra-peritoneal injection of 0.5 mL of BrdU solution (10 mg/mL) into all mice for BrdU-ELISA test, and 0.1 mL of BrdU solution (20 mg/mL) for BrdU-FCM test.
- Skin sensitisation classification
  Validation of skin sensitisation classification was conducted as described in the following decision tree (Figure 1) with the only difference that SI score > 1.6 (BrdU-ELISA method) and SI score > 2.7 (BrdU-Flow Cytometry Method) are considered as positive results. No EC1.6 or EC2.7 were recalculated based on those results. All studies with a non-linear dose-response were not included in the dataset.
- Weight of evidence
  Whenever other *in vivo* and/or *in vitro* studies were available on the same substance (OECD 406, 442C, 442D, 442E), a weight of evidence approach was also considered to strengthen validation of skin sensitisation classification.

4 Defining the algorithm - OECD Principle 2

4.1 Type of model

The model is an expert-based decision tree generating a final outcome result based on the results of four independent sub-models:

- Hapten/prohapten detection (based on MechoA) v1.2 (hereafter referred as Hapt v1.2)
- Autoxidation (prehapten detection) v1.0 (hereafter referred as AutOx v1.0)

- Skin penetration v1.0 (hereafter referred as SkinAbs v1.0)
- Expected LLNA positive LLNA+ v1.1 (hereafter referred as LLNA+ v1.1)

Those sub-models are either based on structural alerts or physicochemical property thresholds.

4.2 Explicit algorithm

The algorithm is an expert-based decision tree generating a final outcome according to the results of each of the four sub-models.

4.3 Descriptors in the model

The model encompasses three types of descriptors:

- mechanistic insight,
- 2D structural fragments (alerts),
- physicochemical properties.

4.4 Descriptor selection

Descriptors for the model were selected based on literature, experimental data and expert judgment.
There was an *a priori* assumption that some chemical fragments associated with a mechanistic interpretation can be used to model skin sensitisation potential. In addition, some of the physicochemical parameters proven to be determinant in the skin absorption process were selected as descriptors: log K_{OW} and Molecular Weight (MW) (R7c chapter ECHA guidance, 2017).

4.5 Algorithm and descriptor generation

The descriptor of the sub-models Hapt v1.2, AutOx v1.0 and LLNA+ v1.1 are structural alerts. These alerts were determined by experts and designed to identify reactive chemical features (haptens), chemical moieties that are bioactivated into reactive species (pro-haptens), and chemical structures that are autoxidized to reactive species (pre-haptens). Hapt v1.2 is able to discriminate if whether the substance is a hapten, a prohapten, or if it does not covalently bind to protein when the sub-model is in its applicability domain. In addition, for LLNA+ v1.1, structural alerts were developed based on chemical features associated with positivity in LLNA tests but not in other *in vivo* tests and are detected as “positive in LLNA only”.
The sub-model SkinAbs v1.0, is based on physico-chemical descriptor thresholds for log K_{OW} and MW. These thresholds were chosen according to Chapter R7c ECHA guidance (ECHA, 2017). According to these thresholds, skin penetration is predicted as “sufficient” or “not sufficient”.
The decision tree of the models was developed based on expert judgment. It is inspired by the skin sensitisation AOP (OECD, 2014).

4.6 Software name and version for descriptor generation

The descriptors (independent variables) and endpoint values (dependent variables) were generated by expert judgement, experimentally derived and retrieved from various literature resources including some publicly disseminated databases as well as some confidential data available within KREATiS internal database. When experimental physicochemical properties were lacking, those were generated by iSafeRat desktop v4.0 and EPIWEB v4.1.

4.7 Chemicals/Descriptors ratio

Chemical/Descriptors ratios are not applicable as the model is based on structural alerts rather than statistically based.

5 Defining the applicability domain - OECD Principle 3

5.1 Description of the applicability domain of the model

a) Fixed or probabilistic boundaries

The following domains (response, descriptor, structural fragment, mechanistic and metabolic domains) are fixed boundaries based structural alerts and on training set data.

b) Response domain

The response domain of the final outcome prediction is defined by four sub-models: Hapt v1.2, AutOx v1.0, LLNA+ v1.1 and SkinAbs v1.0.

iSafeRat® CLASS provides four possible final outcomes:

- Skin sensitiser,
- Not skin sensitiser,
- Expected positive in LLNA only,
- N/D.

Skin sensitiser (*i.e.*, positive), Expected positive in LLNA only (*i.e.*, positive) and Not skin sensitiser (*i.e.*, negative) predictions are considered to be within the applicability domain of iSafeRat® CLASS. N/D outcome indicates that the test substance is outside of the response domain of iSafeRat® CLASS. When the substance is in the applicability domain of iSafeRat® CLASS, all different outcomes of the model are gathered in the following table:

Table 3: Possible response of CLASS model

Hapt v1.2	AutOx v1.0	LLNA+ v1.1	SkinAbs v1.0	Final Outcome of CLASS
Positive	Positive	Yes	Sufficient	**Sensitiser**
Negative	Positive	Yes	Sufficient	**Sensitiser**
Positive	Negative	Yes	Sufficient	**Sensitiser**
Positive	Positive	No	Sufficient	**Sensitiser**
Negative	Positive	No	Sufficient	**Sensitiser**
Positive	Negative	No	Sufficient	**Sensitiser**
Negative	Negative	Yes	Sufficient	**Expected positive in LLNA only**
Positive	Positive	Yes	Insufficient	**Not Sensitiser**
Negative	Positive	Yes	Insufficient	**Not Sensitiser**
Positive	Negative	Yes	Insufficient	**Not Sensitiser**
Positive	Positive	No	Insufficient	**Not Sensitiser**
Positive	Negative	No	Insufficient	**Not Sensitiser**
Negative	Positive	No	Insufficient	**Not Sensitiser**
Negative	Negative	Yes	Insufficient	**Not Sensitiser**
Negative	Negative	No	Sufficient	**Not Sensitiser**
Negative	Negative	No	Insufficient	**Not Sensitiser**

The substance is outside of response domain of iSafeRat® CLASS (N/D) if the substance is outside of the applicability domain of Hapt v1.2 and/or SkinAbs v1.0 sub-models.
The Hapt v1.2 sub-model is outside of its applicability domain in cases where the test substance contains one or more unrecognised structural fragments identified by the software.
The SkinAbs v1.0 sub-model is outside of its applicability domain in cases where log K_{OW} and/or MW cannot be predicted by iSafeRat or log K_{OW} and/or molecular weight is missing as user input.

c) Descriptor domain

The test item should be a mono-constituent organic substance with a defined SMILES. A MW and a predicted or experimental log K_{OW} values are mandatory. Furthermore, the descriptor domain of iSafeRat® CLASS corresponds to its structural and mechanistic domain.
For SkinAbs v1.0 module, all substances with a MW ≤ 500 g/mol or with a $-1 \leq \log K_{OW} \leq 4$ will give a "sufficient skin penetration" outcome based on ECHA R7c (ECHA, 2023). Otherwise, the module will give an "insufficient skin penetration" response.

d) Structural fragment domain

The test item must be a mono-constituent organic substance. If a substance is not recognised by the software, if the substances is an UVCB or a multi-constituent or if the substance is not an organic chemical, then the substance is considered as outside of the structural fragment domain of the model.

The structural fragment domain is primarily defined by the applicability domain of Hapten v1.2.
This sub-model is based on MechoA Premium alerts that are related to the molecular structure of the substance and structural alert exclusion rules. Hence, if a structural alert is detected by MechoA Premium scheme and no exclusion rules have been met by the sub-model, the substance is inside the structural domain of Hapten v1.2.

The following list of chemical moieties provide an insight of the broad structural domain:

- Aldehydes (including benzaldehyde and α,β-unsaturated)
- Alcohol (including α,β-unsaturated)
- Alkane
- Alkenes
- Amines & diamines
- Epoxides
- Ester (including α,β-unsaturated)
- Halogenated hydrocarbons
- Imines
- Ketones (including α,β-unsaturated)
- Quinones and precursors

Note1: The number of times each fragment is present in the test item is not taken into account.

e) Mechanistic domain

iSafeRat® CLASS predicts test substances to be skin sensitisers or non-skin sensitisers in a LLNA study.

A substance is predicted as a “skin sensitiser” in cases where the following mechanisms are detected:

- Hapten: hard-electrophile reactivity (MechoA 3.1), soft-electrophiles reactivity (MechoA 3.2) & spontaneous abiotic radical generation (MechoA 3.3) (based on Hapt v1.2)
- AND/OR: Pro-hapten: pro-reactivity (MechoA 4.3) & red-ox cycling (MechoA 4.4) & ROS generation (MechoA 5.1) (based on Hapt v1.2)
- AND/OR: Pre-hapten: alkenes susceptible to undergo autoxidation (based on AutOx v1.0)
- AND: “sufficient skin penetration” (based on SkinAbs v1.0)

In cases where a test substance does not trigger any structural alerts related to these mechanisms, the test substance is considered to be a non-skin sensitiser in the LLNA study. However, if a substance contains a chemical feature associated with positivity in LLNA, but not in other experimental tests (OECD 406 for example), it will be predicted as “Expected positive in LLNA only”.

A substance is a “non-skin sensitiser” in cases where the substance is not detected to covalently bind to proteins based on the above mechanisms (Hapt v1.2 and AutOx v1.0 are negative, which mean no hapten, prohapten or prehapten was detected) and does not have ”sufficient skin penetration“ to elicit a response (SkinAbs v1.0 result is skin penetration not sufficient) (see section 4.5).

In cases where a test substance is not associated with any molecular initiating event detected by the updated version of the model MechoA Premium scheme or is outside of the applicability domain of MechoA Premium scheme, the test substance is considered out of the applicability domain of iSafeRat® CLASS. (Bauer *et al.*, 2018; Levet *et al*. (soon to be published)).

f) Metabolic domain, if relevant

Metabolism is taken into account in the mechanistic domain within the MechoA scheme uniquely. No other module takes metabolism into account.

g) Possible defined (graphical) expression of how the descriptor values of the chemicals in the training set are distributed in relation to the endpoint values predicted by the model.

Not relevant.

5.2 Method used to assess the applicability domain

First of all, the test substance is verified to be within the mechanistic domain of the model. If every structural fragment is recognized by the MechoA Premium profiler, hence, the substance is considered within the applicability domain of the model. In cases where a test substance contains unrecognised chemical features or cannot be associated with a MechoA, the test substance is considered out of applicability domain. Furthermore, the MW and log K_{OW} of the test substance are required to estimate if skin absorption is sufficient to trigger skin sensitisation. While the MW is automatically calculated by iSafeRat® desktop, log K_{OW} can either be filled in by the user (using an experimental or other predicted value) or predicted by iSafeRat® KOW v2.0 (and later versions). In the case where no log K_{OW} is added by the user and the test item is out of the applicability domain of iSafeRat® KOW v2.0, the test substance is considered out of the applicability domain of iSafeRat® CLASS.

If a positive alert is fired by the Hapt v1.2 and/or AutOx v1.0 sub-models, plus the skin penetration (SkinAbs v1.0) is predicted to be sufficient, the final outcome is "skin sensitiser" (*i.e.*, predicted positive in LLNA). Further details on the mechanism of toxic action are available by clicking on *More Details* in the *MechoA Premium v1.0* window. If negative alerts are generated by the Hapt v1.2 & AutOx v1.0 sub-models, but an alert is fired by the LLNA+ v1.1 sub-model (*i.e.*, predicted positive in LLNA, but expected to be negative in other *in vitro* and *in vivo* tests (e.g. GPMT, Buehler, etc.), and skin penetration (SkinAbs v1.0) is predicted to be sufficient, then the final conclusion in the box "Classification Labelling for Skin Sensitisation CLASS v1.4" will come out as "Expected positive in LLNA only" (*i.e.*, predicted positive in LLNA).

If no alert is triggered and/or skin penetration (SkinAbs v1.0) is insufficient, then the final outcome is "not skin sensitiser" (*i.e.*, predicted negative in LLNA).

In cases where log K_{OW} input is lacking and/or the SMILES cannot be handled by the software, no prediction is generated, and the outcome is "N/D".

In cases where Hapt v1.2 and SkinAbs v1.0 are outside of applicability domain (N/D), the prediction is outside of iSafeRat® CLASS applicability domain.

5.3 Software name and version for applicability domain assessment

The applicability domain assessment was performed using iSafeRat® Desktop v4.3.14.

5.4 Limits of applicability

The limits of the applicability domain are mainly defined by the limits of each structural alerts, the presence of unrecognised chemical features on a test item and missing physicochemical parameters.

6 Defining goodness-of-fit and robustness (internal validation) - OECD principle 4

6.1 Availability of the training set

The training set of the model is proprietary and has not been made publicly available. The training set of the model may be shared with regulatory authorities upon their request.

6.2 Available information for the training set

The following information is not reported as supporting information (see field 9.3) as it is proprietary. All queries must be directly addressed to KREATiS SAS.
CAS RN: YES (confidential business information)
Chemical Name: YES (confidential business information)
SMILES: YES (confidential business information)
Formula: NO
INChI: NO

MOL file: NO

6.3 Data for each descriptor variable for the training set

The training set of the model is proprietary and has not been made publicly available yet. All queries must be directly addressed to KREATiS SAS.

6.4 Data for the dependent variable for the training set

The training set of the model is proprietary and has not been made publicly available yet. All queries must be directly addressed to KREATiS SAS.
Data available: log K_{OW} values, LLNA outcome, data source.

6.5 Other information about the training set

The training set was composed of 467 LLNA data, including 223 negatives and 246 positives. All queries must be directly addressed to KREATiS SAS.

6.6 Pre-processing of data before modelling

Pre-processing was performed based on description provided in section 3.6 and 3.7.

6.7 Statistics for goodness-of-fit

Table 4: Statistics for goodness-of-fit

Dataset	Accuracy	Sensitivity	Precision (positive)	Specificity	Precision (negative)	Balanced accuracy	True positive	True negative	False positive	False negative
Training set (n=467)	78.6%	84%	77,1%	72.6%	80,6%	78.6%	205	162	61	39

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

This is not applicable.

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

This is not applicable.

6.10 Robustness - Statistics obtained by Y-scrambling

This is not applicable.

6.11 Robustness - Statistics obtained by bootstrap

This is not applicable.

6.12 Robustness - Statistics obtained by other methods

None.

7 Defining predictivity (External validation) - OECD Principle 4

7.1 Availability of the external validation set

The external validation set of the model is proprietary and has not been made publicly available. The external validation set of the model can be shared with other regulatory authorities on request.

7.2 Available information for the external validation set

CAS RN: YES
Chemical Name: YES
SMILES: YES
Formula: NO

INChI: NO
MOL file: NO

7.3 Data for each descriptor variable for the external validation set

The external validation set of the model is proprietary and has not been made publicly available yet

7.4 Data for the dependent variable for the external validation set

The external validation set of the model is proprietary and has not been made publicly available yet.
Data available: log K_{OW} values, LLNA outcome, data source.

7.5 Other information about the external validation set

The external validation set was composed of 123 LLNA data, including 46 negatives and 77 positives.

7.6 Experimental design of test set

The data splitting workflow applies a modular scheme of grouping entries of the reference dataset in order to split into training and test subsets. The splitting stages allowed us to balance the training and test subsets with respect to the functional classes and chemical structures. Furthermore, by varying the sequential order of splitting stages it was possible to refine the resulting dataset splitting, and hence to split the dataset in the most suitable way for learning from the training subset and to explore the predictive performance of the model.

Four levels of grouping/splitting are defined:

- "Level 0" is termed as the grouping of entries by preset criteria. Within the CLASS® modelling process, Level 0 splitting refers to grouping of entries by MechoA classification (mechanistic grouping), allowing the subsets to be balanced in terms of mechanistic properties of constituents (the mechanistic properties being inaccessible by simple structural description without implementation of knowledge of biological effect associated with different chemical structure elements).
- "Level 1" is termed as the grouping of dataset entries by structural similarity measured by Tanimoto index (Tc) based upon description of the dataset entries by a molecular fingerprint. Within CLASS®, a KREATiS-modified version of the PubChem molecular fingerprint adapted for real-value counted features was used. The Level 1 implements a structural grouping and allowed us to balance the training and test subsets in terms of chemical space representatives.
- "Level 2" is termed as the grouping of dataset entries by their experimental functional classification, *i.e.,* Positives (sensitiser) and Negatives (not sensitiser) in respect to the endpoint of interest.
- "Level 3" is termed as the splitting of dataset entries into training and test subsets. The splitting of the clusters obtained by the previously applied grouping stages is performed using the Kennard-Stone algorithm operating upon the pairwise molecular distances defined as (1-Tc).
- Furthermore, an additional postprocessing step is applied on singleton clusters within both branches defined by the functional classification (Level 2) grouping. Hence, 10% of the Positive and 10% of the Negative singleton clusters were assigned to the test subset. As singleton clusters are further apart in the structural similarity space from substances included in multi-member clusters, the postprocessing enriches the test subset with substances that allow challenges of the applicability domain defined by the training subset.

The splitting scenario was set as the following order of stages: Level 0 → Level 1 → Level 2 → Level 3.
The ratio of substance between train set and test set was 0.79/0.21 among the 590 molecules in the dataset.

7.7 Predictivity - Statistics obtained by external validation

Based on this external validation set, performances were:

Table 5: Statistics for the external validation set

Dataset	Accuracy	Sensitivity	Precision (positive)	Specificity	Precision (negative)	Balanced accuracy	True positive	True negative	False positive	False negative
External validation set (n=123)	75.5%	85.7%	77,6%	58.7%	71.1%	72.6%	66	27	19	11

Accuracy (*i.e.*, percentage of accurate predictions of model)
Sensitivity (*i.e.*, percentage of sensitisers that are predicted as such by a model)
Precision (positive) (*i.e.*, percentage of correctly predicted positives among the predicted positives).
Specificity (*i.e.*, percentage of non-sensitisers predicted as such by a model).
Precision (negative) (*i.e.*, percentage of correctly predicted negatives among the predicted negatives).
Balanced accuracy (*i.e.*, [sensitivity + specificity]/2)

7.8 Predictivity - Assessment of the external validation set

From the statistics obtained from the external validation set (test set), it is possible to conclude that the validation data of CLASS performed well for a large diversity of structures. Variations in performance were observed between the training and the external validation sets.
The accuracy, balanced accuracy and the specificity are slightly lower in the test set, while the sensitivity was 1.7% higher than the training set.
Overall, the model has a good predictive capacity to predict positive skin sensitiser because the test has a high sensitivity 85.7%. However, achieving a specificity of 58.7% means that the model accurately identifies only 58.7% of substances as non-sensitisers, while incorrectly classifying 41.3% of non-sensitiser substances as sensitisers. This can be viewed as a worst-case scenario which mean that the model is considered to be conservative.

7.9 Comments on the external validation of the model

None.

8 Providing a mechanistic interpretation - OECD Principle 5

8.1 Mechanistic basis of the model

All structural alerts are established based on scientific literature and experimental data. In order to ease the expertise of predictions by the user, the outcome of the four profilers is given by the model, in addition to the final outcome. Furthermore, the details of the mechanism of toxic action (MechoA) are described in the MechoA Premium tab.

8.2 A priori or a posteriori mechanistic interpretation

The mechanistic basis of the model was developed *a priori* by examining the toxicological and mechanistic evidence before developing the structure-activity relationship sub-models. Determining the $\log K_{ow}$ and MW threshold was done *a posteriori*.

8.3 Other information about the mechanistic interpretation

To allow a better understanding of the methodology, no inexplicable molecular descriptors or modelling algorithms were included. The mechanistic basis of the alerts is detailed and available within MechoA Premium.

9 Miscellaneous information

9.1 Comments

This QMRF can be used as a reference document for QPRF reports providing iSafeRat® CLASS v1.4 predictions.

9.2 Bibliography

Bauer FJ, Thomas PC, Fouchard SY, Neunlist SJM. *High-accuracy prediction of mechanisms of action using structural alerts*. Comput Toxicol. 2018;7:36-45.

Butina D. *Unsupervised data base clustering based on daylight's fingerprint and tanimoto similarity: A fast and automated way to cluster small and large data sets*. J Chem Inf Comput. 1999;39(4):747-750.

European Chemicals Agency. *Guidance on information requirements and chemical safety assessment: chapter R.7c: Endpoint specific guidance*. Helsinki: European Chemicals Agency. 2023.

Karlberg AT, Bergström MA, Börje A, Luthman K, Nilsson JL. *Allergic contact dermatitis-formation, structural requirements, and reactivity of skin sensitizers*. Chem Res Toxicol. 2008;21(1):53-69.

Kennard RW, Stone LA. *Computer aided design of experiments*. Technometrics. 1969;11(1):137-148.

Levet G, Darracq-Ghitalla-Ciock M, Bourgart E, Bauer FJ, Ay-Albrecht E, Todorov Z, Thomas PC, *et al*. *Classification & labelling assessment for skin sensitisation (CLASS) update: a mechanistic in silico model to predict skin sensitisation potential*. Eurotox. 2024.

Levet G *et al*. [*Publication on an updated version of MechoA scheme* – soon to be published].

Organisation for Economic Co-operation and Development. *Test No. 429: Skin Sensitisation: Local Lymph Node Assay*. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. 2010a.

Organisation for Economic Co-operation and Development. *Test No. 442A: Skin Sensitization: Local Lymph Node Assay: DA*. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. 2010b.

Organisation for Economic Co-operation and Development. *The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins*. OECD Series on Testing and Assessment, OECD Publishing, Paris. 2014;168.

Organisation for Economic Co-operation and Development. *Test No. 442B: Skin Sensitization: Local Lymph Node Assay: BrdU-ELISA or –FCM*. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. 2018.

Ryan CA, Chaney JG, Gerberick GF, Kern PS, Dearman RJ, Kimber I, Basketter DA. *Extrapolating local lymph node assay EC3 values to estimate relative sensitizing potency*. Cutan Ocul Toxicol. 2007;26(2):135-45.

9.3 Supporting information

Training set(s):
Proprietary. All the queries must be directly addressed to KREATiS SAS.

Test set(s):
Proprietary. All the queries must be directly addressed to KREATiS SAS.

Supporting information:
None.

10 Summary (KREATiS QMRF Database)

10.1 QMRF number (For KREATiS internal records only)

KTS/QMRF/CLASSv1.4/05

10.2 Publication date

23 April 2025

10.3 Keywords

iSafeRat®; Skin sensitisation, potential

10.4 Comments

None.