Classification & Labelling Assessment for Skin Sensitisation (CLASS): AN ADVANCED MULTI-SAR MODEL TO PREDICT SKIN SENSITISATION

G. Levet[1], M. Darracq-Ghitalla-Ciock[1], E. Bourgart[1], F. J. Bauer[1], E. Ay-Albrecht[1], Z. Todorov[1], P. C. Thomas[1], C. Charmeau-Genevois[1]

[1]KREATiS, 38080 L'Isle d'Abeau – France

Contact us : www.kreatis.eu / contact@kreatis.eu

Introduction

iSafeRat® CLASS (Classification & Labelling Assessment for Skin Sensitisation) is designed to **predict skin sensitisation** (SS), otherwise known as allergic contact dermatitis (ACD) of organic chemical substances. SS is predicted through four sub-models: skin penetration (SkinAbs v1.0), hapten (i.e. protein adduct formation) & pro-hapten detection (Hapt v1.2), pre-hapten detection through autoxidation (Autox v1.0) and expected positivity in the LLNA test only (LLNA+ v1.1). To build this model, a **splitting algorithm** was created internally to ensures that the model's performance is evaluated fairly and representatively. The first version which predicted the potential for positive or negative sensitisation was developed in 2023. Since then, major improvements have been integrated which **increase confidence in predictions and broaden the applicability domain**. In the meantime, regulatory needs were redefined in the recent publication of the QSAR Assessment Framework (QAF)[1]. This framework was designed to help regulators across the OECD member states to **better assess the prediction capacity of QSARs** by using a reflective check list scheme to arrive at a consensus result. QAF sets **4 new principles** for prediction validity, in addition to the **5 OECD principles for model validity**. This poster presents the improvements of CLASS for SS potential prediction and the related updates in **compliance with the QAF requirements**.

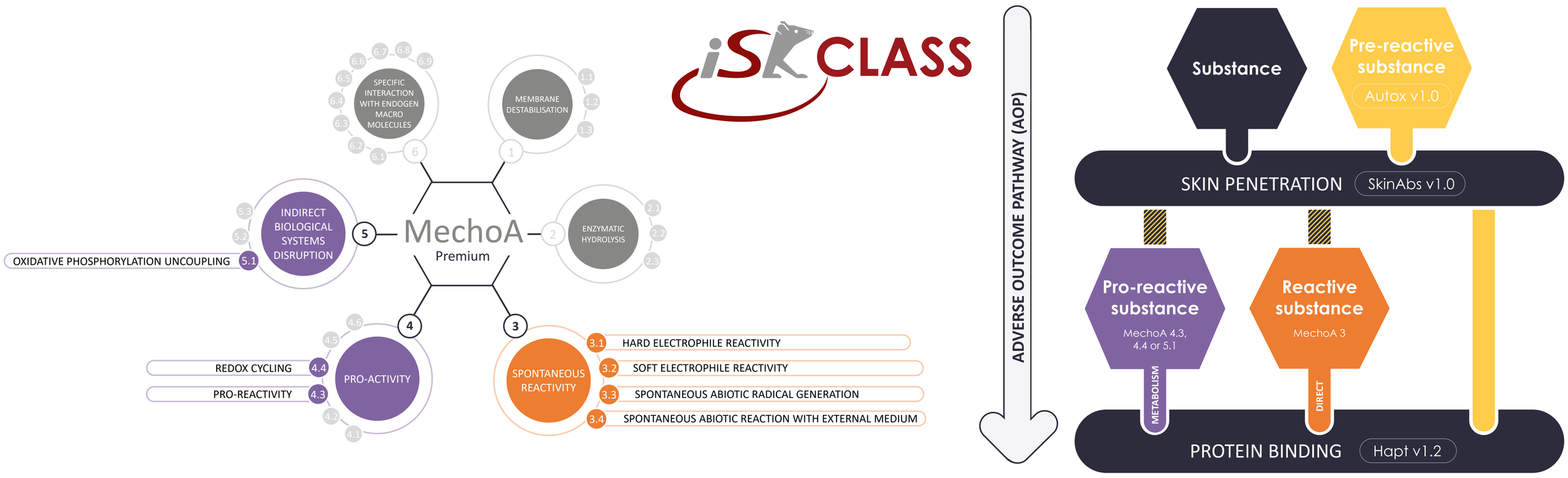

Figure 1. *Classification & Labelling Assessment for Skin Sensitisation (CLASS) model is combining 4 sub-models (Autox v1.0, Skin Abs v1.0, Hapt v1.2 and LLNA+ v1.1 which is not represented here) to assess skin sensitisation through MechoA Premium.*

Methods

Database: LLNA data from NiceATM, CosUE, ECHA dossiers and proprietary data were evaluated for their reliability according to the OECD 429, 442a and 442b guideline requirements, to select only high-quality data. Inorganic substances and multi-constituents were removed from the database. The new database is composed of validated data on 590 substances, which are **structurally heterogeneous**. Besides, the dataset is **balanced** between number of positives and negatives (321 sensitisers and 269 non-sensitisers).

Improvement of the model reliability & applicability domain: A new splitting method, based on similarity (structural and mechanistic) and Kennard-Stone algorithm operating upon the pairwise molecular distances, was applied. We (re)defined the structural alerts and exclusion rules for the sub-models based on the addition of new substances in the training set and further literature searches. Furthermore, to distinguish positive and negative sensitisers, an additional fragmentation approach was implemented.

Improving expert knowledge confidence in the prediction: an automated analogue search based on mechanistic and structural information was designed.

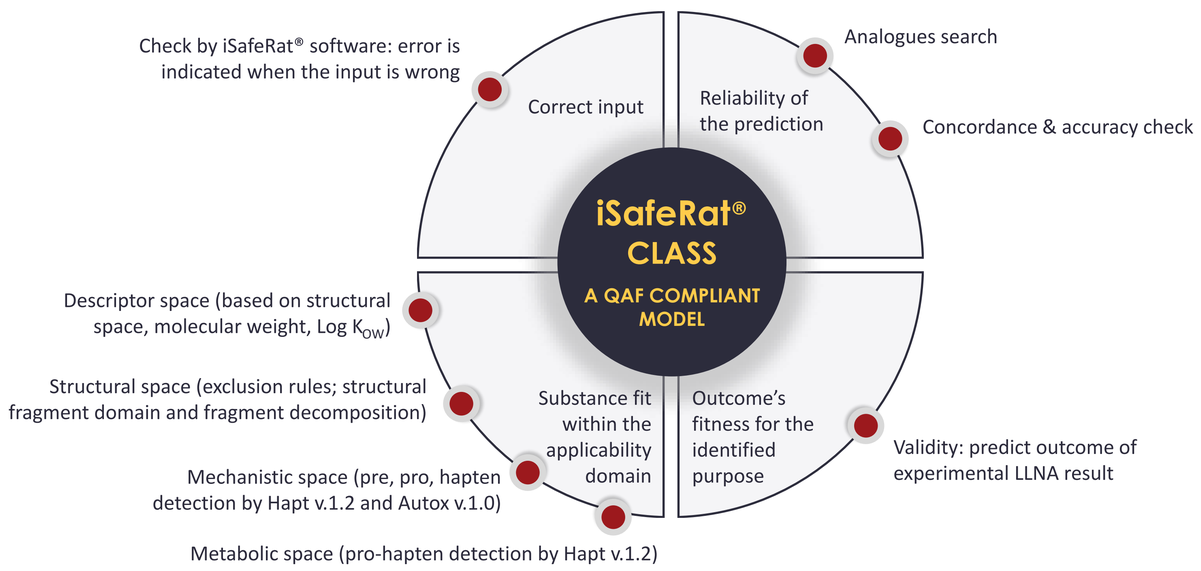

Figure 2. *QAF compliancy of CLASS*

Results and discussion

The update of the model improved the QAF compliancy and gives greater insight into the confidence of the predictions using iSafeRat® CLASS. Statistics of the v1.4 of the model are available in **Table 1**. based on structural and mechanistic information. In order to fit QAF requirements[1], the model was developed to comply with the 4 principles, as described in **Figure 2**.

Dataset	Size (N)	Precision (%)	Accuracy (%)	Balanced accuracy (%)	Sensitivity (%)	Specificity (%)
Training	467	77.1	78.6	78.3	**84.0**	**72.6**
Test	123	77.6	75.6	72.2	**85.7**	**58.7**

Table 1. *Statistics of iSafeRat® CLASS v1.4 implemented in iSafeRat® Desktop v4.3.14.*

The splitting method was designed to keep structural and mechanistic heterogeneity between the training and test sets. The low specificity observed for the test set could be explained because of the negative data which are underrepresented (both structurally or mechanistically). To fill this gap, the authors aim to increase the number of negative data to ensure a better representativity between the train and test sets.

Furthermore, an additional fragmentation approach was implemented to improves the automatic structural domain evaluation for both positive and negative predictions given by the model. This reinforces and maximises the reliability of the predictions to fit QAF requirements[1].

In addition, to better fit QAF requirements[1], analogues are searched based on both structure similarity and MechoA[2]. Analogue searches are therefore not based exclusively on mathematical and chemical methodology, but also through toxicological similarity. Furthermore, predictions and experimental results of the **analogues are provided in order to assess the reliability of the prediction**. This expert knowledge tool gives more insight into the prediction decision and can therefore be used as a way to validate the relevance of analogues located in the training set of the model for both negative and positive predictions, as requested by the QAF[1].

Conclusion

CLASS provides prediction for skin sensitisation potential with a good reliability in general. With the latest work a **better insight into the applicability domain** is now given through a **fragmentation approach methodology**. Furthermore, the analogue search helps the user **to find relevant analogues** in the dataset in order to expertise the reliability of a prediction. The model is **fully compliant with ECHA requirements according to the QAF.** Therefore, it better fills in the QPRF regulatory document requested in the case of an upcoming registration of a substance. Furthermore, sensitisation potency through EC3 prediction will be implemented in the next version of CLASS. The analogue search will also be improved, with EC3 predictions for the analogues identified by the software in order to assess the reliability of the EC3 prediction. Finally, iSafeRat® CLASS is now included in the "me too" process for inclusion in the next version of the OECD guidance of Defined Approaches for Skin Sensitisation (OECD 497).

Scan QR code to download the poster
Available on https://kreatis.eu

[1] OECD. *(Q)SAR assessment framework: Guidance for the regulatory assessment of (quantitative) structure activity relationship models and predictions*. Paris: OECD Publishing. 2023;386.
[2] Bauer FJ *et al. . A new classification algorithm based on mechanisms of action.* Comput Toxicol. 2018;5:8–15.