# AOrTA: a reliable *in silico* tool to predict Acute Oral Toxicity for regulatory purposes

Marie Darracq--Ghitalla-Ciock[1], Emel Ay-Albrecht[1], Carole Charmeau-Genevois[1]

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

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

## Introduction

In the context of the 3Rs, Acute Oral Toxicity (AOT) is one of the last endpoints of REACH Annex VII which needs to be replaced by new alternative methods (NAMs). **Existing *in silico* tools are struggling** to provide models with good performance and reliable prediction in regards of QSAR Assessment Framework[1] for **regulatory use**.

In this context, a new *in silico* global model named AOrTA v1.0 (for "Acute Oral Toxicity Alert"), was developed to **predict Acute Toxicity Estimates** (ATE) for AOT according to the **CLP classification for regulatory purpose**.

## Methods

The dataset of the model has been constituted from AOT studies extracted from ECHA dossiers and scientific literature. Only organic mono-constituents were used to develop the model. Studies were **validated individually and manually** in agreement with the OECD guidelines in order to build a model using a **high-quality dataset**. Dataset splitting was performed using a **genetic algorithm**, optimising parameters to **ensure structural representativity** between the training and test sets. Then, a Support Vector Machine (SVM) was used, with hyperparameters optimised through k-fold cross-validation combined with a genetic algorithm. An applicability domain based on the chemical descriptors used in the model (QFA approach) was also defined. This approach **verifies whether a substance lies within the descriptor space of the model and provides a reliability index** accordingly. In this way, **toxicologists are given an applicability domain index to better assess the confidence of the prediction**. For transparency, **validated $LD_{50}$ values are also reported** for all structural analogues as additional information.

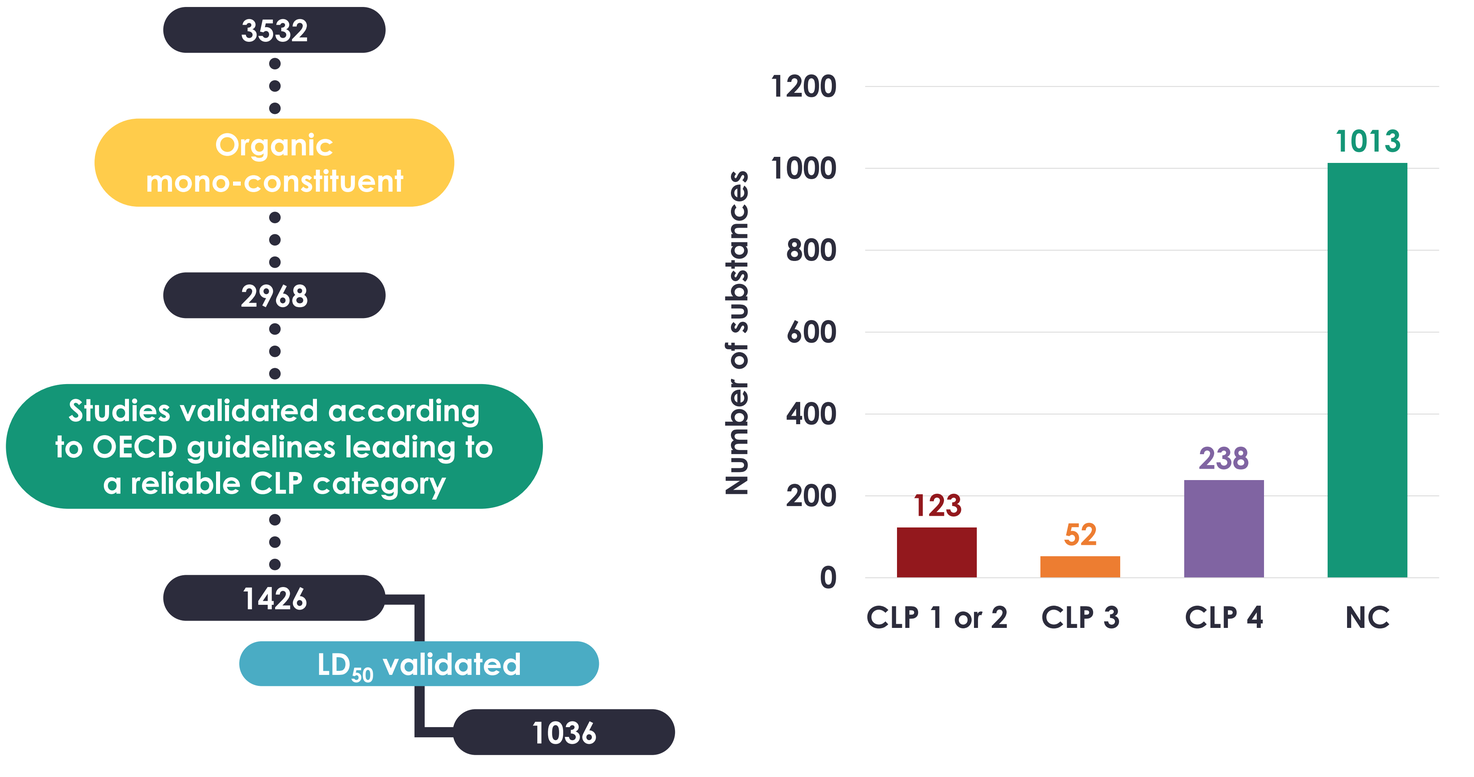


***Figure 1.*** *Manual curation of the rat acute oral toxicity data led to the selection of 1426 data within each CLP category (NC, non-categorised according to CLP).*

## Results

The dataset is **imbalanced due to the lack of data on toxic substances in public databases**, particularly for CLP category 3. Nevertheless, the included studies have been **rigorously expert-reviewed**, ensuring **high data quality**. The developed binary model demonstrated **good performance** in this first version, especially for the "NC" group (Table 1). The model's robustness was evaluated using k-fold cross-validation, showing **limited variability of 5–7%** in the statistics (based on the macro F1-score) across the training set, the training k-folds, and the validation k-folds. These results indicate the **model's stability and its reliable ability to generalise new data**. The **defined applicability domain will further enhanced prediction reliability**.

| | Concordance or accuracy | Sensitivity (categorised) | Precision (categorised) | F1-score (categorised) | Specificity (NC) | Precision (NC) | F1-score (NC) | Macro F1-score |
|---|---|---|---|---|---|---|---|---|
| **Training set (goodness-of-fit)** | 84% | 82% | 68% | 74% | 84% | 92% | 88% | 81% |

***Table 1.*** *Performance metrics for the training set for categorised substances and non-categorised (NC) substances according to CLP.*

| | Mean Macro F1-score | Std. Macro F1-score |
|---|---|---|
| **5-fold train** | 80% | 0.03 |
| **5-fold validation** | 73% | 0.02 |

***Table 2.*** *Results of 5-fold cross-validation for the training and validation sets.*

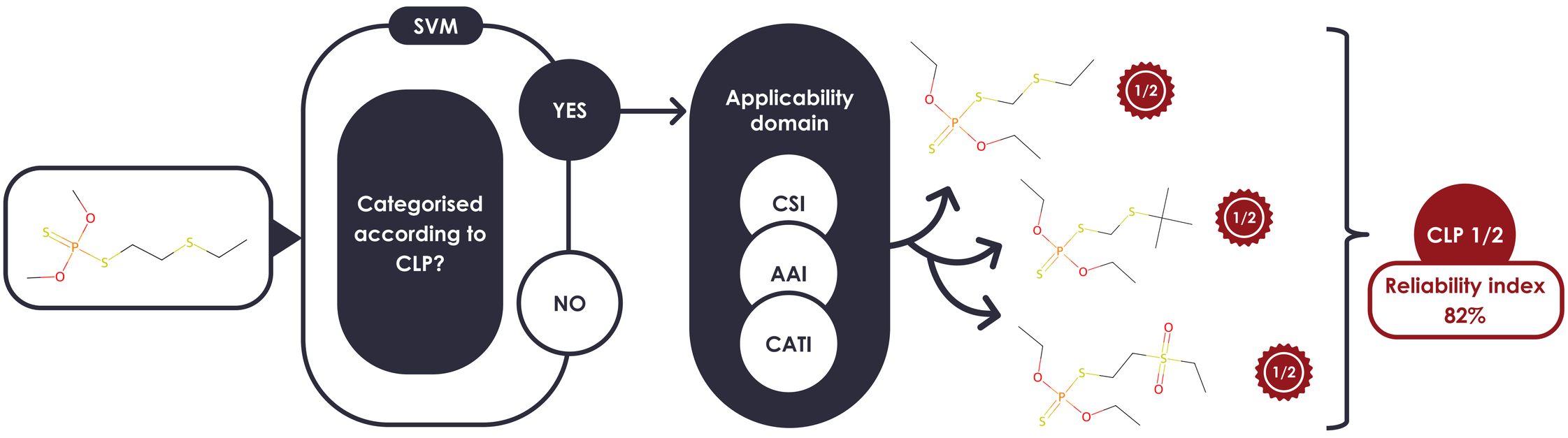


***Figure 2.*** *Example of the prediction of a substance which is not within the dataset and experimentally categorised cat 2 according to NITE Japan-GHS[2]. The substance goes through a binary SVM and, to know if the test item is within the applicability domain, CSI (combined similarity index), AAI (analogue based accuracy index) and CATI (concordance analogue & test item index) are calculated leading to a reliability index of 82% for the prediction of the CLP category.*

## Conclusion

AOrTA v1.0 was developed to assess AOT based on structural features of the substances to **predict ATE with high confidence**. The tool provides sufficient information for **predictions to be evaluated by toxicologists**, supporting the **replacement and refinement of animal testing**. Future improvements to the first version of the model are planned, and **additional local models will be developed** for substances that are not well covered by the global model, in order to enhance the confidence and reliability of AOT predictions.

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

[1] OECD (2024), (Q)SAR Assessment Framework: Guidance for the regulatory assessment of (Quantitative) Structure Activity Relationship models and predictions, Second Edition, OECD Series on Testing and Assessment, No. 405, OECD Publishing, Paris.
[2] NITE Japan-GHS, Latest GHS Classification Results by the Japanese Government for CAS 640-15-3. Available at: https://www.chem-info.nite.go.jp/chem/english/ghs/m-nite-640-15-3e.html.