Screening approach using open access *in silico* models to predict systemic toxicity

Andreas Weber*; **Alexander Molter**; Paul Benndorf; Simone Hoffmann-Doerr

Henkel AG & Co. KGaA, Corporate Toxicology, Duesseldorf, Germany; *Contact: andreas.weber@henkel.com

Initial thoughts

- 70 % of chemicals on the market lack sufficient data for C&L (EEA, 2019)
- More information to address product safety is needed
- *In silico* tools can help to use existing data more efficiently without additional animal data generation
- Current practice: No data, No C&L

➔ System to evaluate these chemicals needs to be cost-effective, simple to use and open-access

Cramer Classes

Cramer Classes Toxic hazard classification by Cramer (extended)_v.2.5)] were used to categorize substances with negative C and M prediction with high confidence (Figure 1)

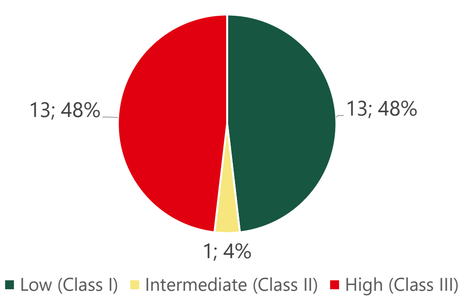

Figure 1: Cramer Classification for C+M positive predicted test items.

Physico-chemical (phys.-chem.) parameters

Cramer Class II substances were categorized as „low" if one of the defined phys.-chem. parameters were met (Figure 2)

Log P: -1 < x < 4
MW: x < 1000 Da
Water solubility: x < 1 mg/l

Figure 2: Defined phys.-chem. Parameters.

Used QSAR models to predict phsy.-chem. Parameters:

- Octanol-water coefficient (log P): Meylan-Kowwin
- Water solubility: IRFMN

Addressing uncertainties and challenges

<u>Applicability domain</u>

Performance of the QSAR models rely on similar chemicals in your training set compared to the target chemical

➔ Only predictions with high certainty were accepted for C+M predictions

➔ Other predictions were termed as "low reliability", and different levels of concerns were assigned (see Table 1)

<u>"Error substances"</u>

Some substances (mainly substances with ionization or containing specific elements like mercury/cadmium) produce errors in the predictions

➔ These substances could not be assessed

Conclusion

An assessment scheme based on open source *in silico* models was used to assign levels of concern to chemicals

Assessment has to be expanded integrating QSAR models covering further endpoints and higher performance as well as integrating PBPK models

Approach

Open access QSAR platforms were compared for availability and performance (e.g., accuracy) of endpoints

➔ QSAR platform **VEGA** gives a margin of confidence/AD in predicted results

QSAR models of VEGA were chosen based on their ability to detect systemic endpoints evaluating performance and comprehensiveness of the training set

➔ Only Mutagenicity (M) and Carcinogenicity (C) models were appropriate

<u>For Mutagenicity:</u> Consensus model v. 1.0.4, which combines results from multiple Mutagenicity QSAR models

<u>For Carcinogenicity:</u> Caesar v. 2.1.10 and ISS v. 1.0.3, which predict Carcinogenicity with higher accuracy compared to the other two models

Predicting Mutagenicity and Carcinogenicity

Mutagenicity and carcinogenicity predictions were combined while also integrating predictions with low confidence (Table 1).

Table 1: A level of concern was assigned based on the predicted C+M properties of a substance, also including the uncertainty of the prediction. C: Carcinogenicity, M: Mutagenicity, L: Low confidence

Mutagenicity	Carcinogenicity	Outcome	Level of concern
Positive	Positive	CM-HIT	**High**
Positive	Low reliability	LC-M-HIT	
Low reliability	Positive	C-LM-HIT	
Negative	Low reliability	LC-M-NOHIT	**Medium**
Low reliability	Negative	C-LM-NOHIT	
Negative	Negative	CM-NOHIT	**↳ Cramer Classes**
Low reliability	Low reliability	Prediction with low confidence	**Not classifiable**

Assessment scheme

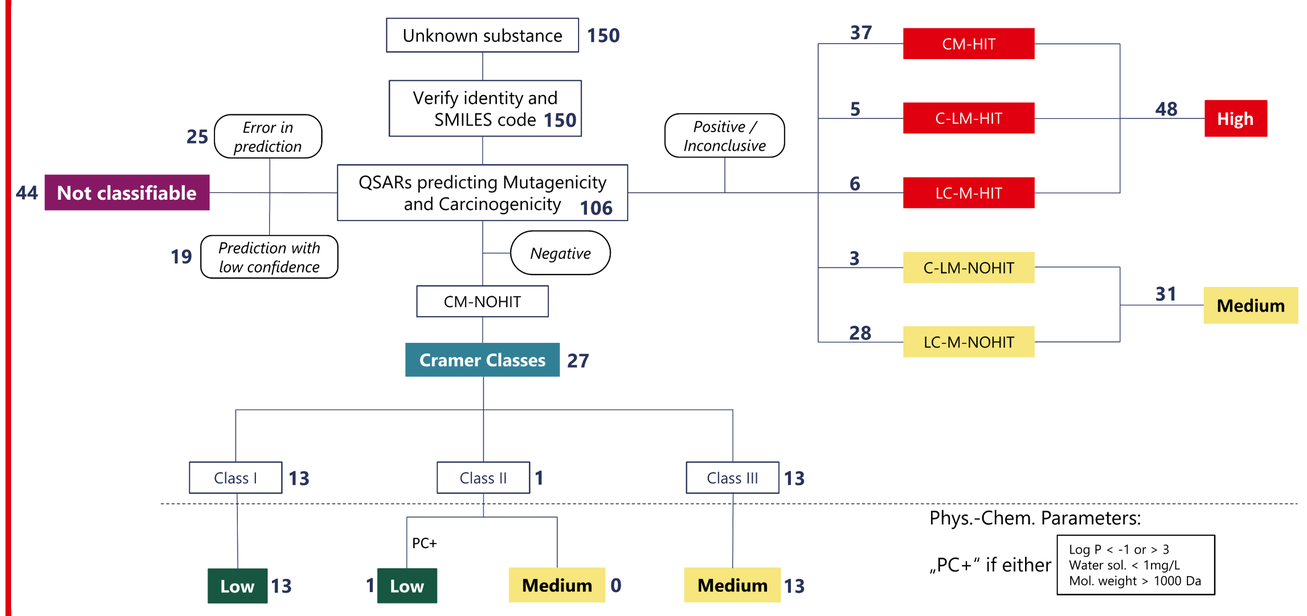

Figure 3: Assessment scheme using C+M QSAR models, Cramer Class prediction and phys.-chem. parameter QSAR models from VEGA. Numbers in **blue** represent the number of assigned substances.

The assessment scheme was used to assign a level of concern to 150 chemicals (see Figure 3, 4):

- 44 substances were not classified based on errors or low confidence in prediction
- 14 substances were assigned to „low"
- 44 substances were assigned to „medium"
- 48 substances were assigned to „high"

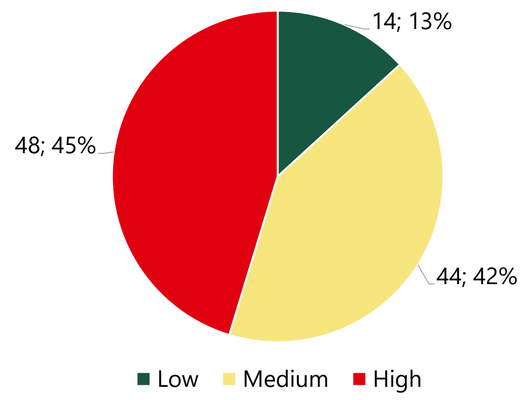

Figure 4: Distribution of assigned levels of concern for the test item set.

Literature

European Environment Agency (EEA): The European environment — state and outlook 2020. Knowledge for transition to a sustainable Europe. (2019) 236 – 240.
G.M. Cramer, R.A. Ford, R.L. Hall. (1976). Estimation of toxic hazard - A decision tree approach. *Food and Cosmetics Toxicology*, 16 (3), 255-276.
VEGA-QSAR: AI inside a platform for predictive toxicology. Proceedings of the workshop "Popularize Artificial Intelligence 2013". Benfenati E, Manganaro A, Gini G., December 5th 2013, Turin, Italy Published on CEUR Workshop Proceedings Vol-1107