

# Introduction to QSARINS-Chem ECO.44 and IVBP-Suite software for the prediction of *in vitro* intrinsic hepatic clearance in human, rat and mouse

QSAR Research Unit in Environmental Chemistry and Ecotoxicology

DiSTA University of Insubria – Varese (Italy)

[ester.papa@uninsubria.it](mailto:ester.papa@uninsubria.it)

[nicola.chirico@uninsubria.it](mailto:nicola.chirico@uninsubria.it)

October 2021



QSAR Research Unit in Environmental Chemistry and Ecotoxicology  
Department of Theoretical and Applied Sciences (DiSTA)  
University of Insubria

## Information

The software and models were developed by:

QSAR Research Unit in Environmental Chemistry and Ecotoxicology  
Department of Theoretical and Applied Sciences (DiSTA)  
University of Insubria, Varese, Italy  
<https://dunant.dista.uninsubria.it/qsar/>

Data collection and curation was performed by ARC Arnot Research & Consulting  
<https://arnotresearch.com>

Funding information: European Chemical Industry Council (CEFIC), grant number: CEFIC-LRI ECO.44 (2018-2020) and extension ECO.44.2 (2021).

We also acknowledge the PhD Course in Chemical and Environmental Sciences (DISCA-University of Insubria) for a PhD fellowship granted to Linda Bertato.

### Contacts

Prof. Ester Papa - e-mail: [ester.papa@uninsubria.it](mailto:ester.papa@uninsubria.it) (project responsible)

Dr. Nicola Chirico, PhD - e-mail: [nicola.chirico@uninsubria.it](mailto:nicola.chirico@uninsubria.it) (software development)

## How to cite:

Please cite IVBP-Suite beta version and QSARINS-Chem ECO.44 beta version in your publications as:

Chirico N., Bertato L., Papa E., In Vitro Biotransformation Prediction-Suite (IVBP-Suite) beta version, 2021, freely downloadable at: <https://dunant.dista.uninsubria.it/qsar/>

Chirico N., Bertato L., Casartelli I., Papa E., QSARINS-Chem ECO.44 beta version, 2021, freely downloadable at: <https://dunant.dista.uninsubria.it/qsar/>

## Limitations of liability and disclaimer of warranty

IVBP-Suite beta version, QSARINS-Chem ECO.44 beta version, and the accompanying materials and manuals are provided "as they are" without warranty of any kind. The authors do not warrant, guarantee, or make any representations, either expressed or implied, regarding the use, or the results from the use of IVBP-Suite beta version and QSARINS-Chem ECO.44 beta version, the accompanying materials and manuals, in terms of correctness, accuracy, reliability, currentness, or otherwise.

You assume the entire risk as to the result and performance of IVBP-Suite beta version and QSARINS-Chem ECO.44 beta version.

In no event shall the authors be liable for any claim, damages or other liability, whether in an action of contract, tort or otherwise, arising from, out of or in connection with IVBP-Suite beta version or the use or other dealings in IVBP-Suite beta version and QSARINS-Chem ECO.44 beta version, even if the authors have been advised of the possibility of such damages.

The IVBP-Suite beta version and QSARINS-Chem ECO.44 beta version software, the accompanying materials and manuals are protected by copyright: 2021, University of Insubria, <http://www.uninsubria.it> - Varese, Italy.

## Introduction

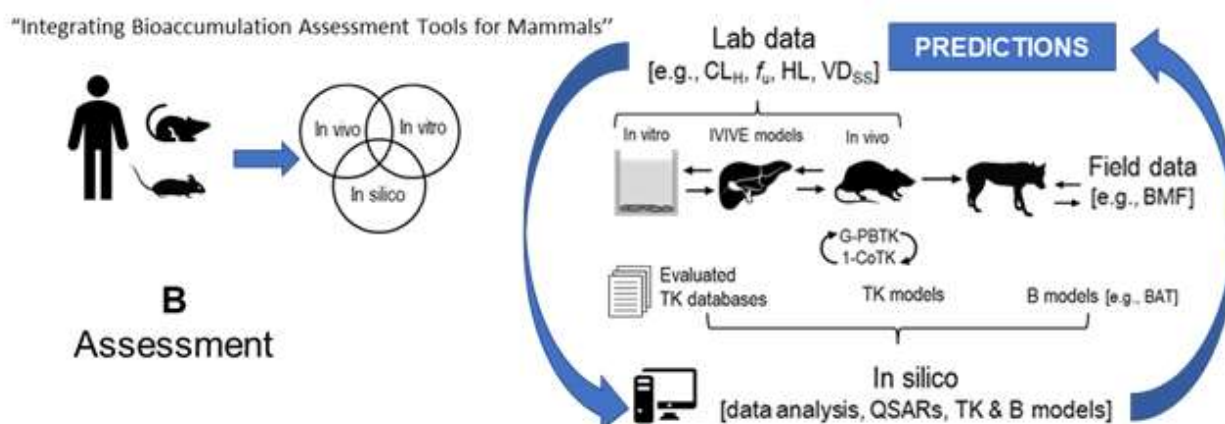
The software QSARINS-Chem ECO.44 and In Vitro Biotransformation Prediction – Suite (IVBP-Suite) have been developed within the context of the CEFIC LRI ECO.44 project “Integrating Bioaccumulation Assessment Tools for Mammals” financed by CEFIC-LRI between 2018 and 2020 and the subsequent extension ECO.44.2 (2021).

The aim of the ECO.44 project was to develop a toxicokinetic framework for the assessment of bioaccumulation in mammals by integrating different data streams and quantitative approaches.

In 2021 the project was extended to create dedicated software to support the application of the over 100 QSAR models developed during the ECO.44 project.

The role of the QSAR Research Unit in Environmental Chemistry and Ecotoxicology (University of Insubria – Varese, Italy) was mainly focused on the development of QSAR models for the prediction of the *in vitro* intrinsic hepatic clearance in rat, mouse and human using data that were collected and curated during the ECO.44 project by our partners at Arnot Research and Consulting - ARC (Toronto, Canada).

## The CEFIC-LRI ECO.44 project



## The *in vitro* intrinsic hepatic clearance database

The database used in this study counts thousands of *in vitro* hepatic clearance values measured in human, rats and mouse for organic molecules characterized by heterogeneous structures. Most of the compiled data have been measured in microsomes or hepatocytes and only minor part have been tested in S9.

The *in vitro* data used to develop our QSARs are available in the EAS-E-Suite online database developed and released by ARC (<https://arnotresearch.com/eas-e-suite/>)

Models developed on S9 data are not included in QSARINS-Chem ECO.44 and IVBP-Suite software.

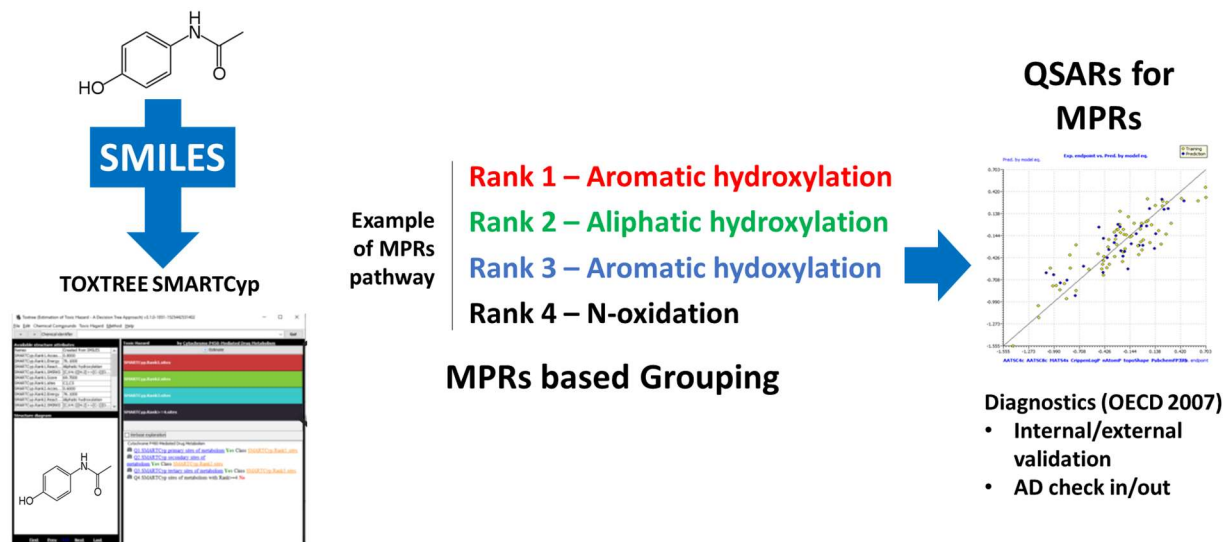
## Reactivity-based modeling strategy

A QSAR modelling strategy based on similarity defined by metabolic reactivity was used to group chemicals and to generate the ECO.44 QSARs developed for human and rodents *in vitro* metabolism, tested in hepatocytes and microsomes. To this end, the SMARTCyp module embedded in Toxtree software (Toxtree v. 3.1.0; Patlewicz G. et al. SAR QSAR Environ Res. 2008; Rydberg, P. et al. ACS Med. Chem. Lett. 2010; Rydberg, P. et al. Bioinformatics 2010) was applied before the development of the QSARs to group chemicals hierarchically according to ranks identified by the software. The SMARTCyp module ranks in a chemical structure the most likely sites of reaction mediated by cytochrome P450, for a total of 15 metabolic reactions (Toxtree v. 3.10). On this basis, starting from SMILES encoding for the molecular structures available in the studied datasets, up to 4 sites responding to 9 possible CYP-mediated reactions (aromatic hydroxylation, N-dealkylation, amine hydroxylation, O-dealkylation, aliphatic hydroxylation, N-oxidation, S-oxidation, epoxidation and alcohol oxidation) were ranked for each chemical. Chemicals were then grouped according to their common reactivity pathway.

Models were generated hierarchically moving from larger groups of chemicals (whose similarity was evaluated by means of the most probable reaction - MPR - identified at the first rank) to smaller groups, whose similarity was evaluated by means of sequential partitioning of the chemicals based on MPRs identified at the first, second, third or fourth rank.

The criteria used to move from one rank to the next one was based on the quality of the QSAR (i.e.,  $R^2 > 0.6$  and  $Q^2 > 0.5$ ) developed for each specific rank. These QSARs, based on reactivity pathways identified by the SMARTCyp module, are mostly associated to a single reaction detected as the most probable in a specific rank, among the nine listed above. However, in some cases, the SMARTCyp module identifies multiple MPRs for a specific rank. For this reason, some QSARs have been defined as multi-reaction, since they were developed for groups of chemicals which share the same pathway of MPRs identified by the SMARTCyp module for a specific rank.

## Reactivity-based grouping and modelling approach



## QSARINS-Chem ECO.44 and IVBP-Suite

In this project more than 100 QSAR models were developed, therefore two software, QSARINS-Chem ECO.44 and IVBP-Suite, were created to simplify their application. QMRF reports were compiled and included in the software to support the regulatory application of these QSARs.

- **QSARINS-Chem ECO.44** is a simplified version of the software QSARINS-Chem standalone version (Chirico N. et al. J. Comput. Chem. 2021) which is completely dedicated to the storage and the application of ECO.44 QSAR models. This software provides a transparent overview of the statistical properties of each ECO.44 QSAR as well as of the training sets used to develop the models. It is suitable for detailed analysis of the predictions calculated by the single models as well as to evaluate their inclusion in the applicability domain of the original model. However, the application of QSARINS-Chem ECO.44 does not assist the user in the choice of the correct model to be applied to new chemicals. This information should be known by the user or alternatively should be generated by profiling manually the chemical of interest using Toxtree with the SMARTCyp module, to identify the reactivity pathway and the related reactions. Knowing this information, is then possible to select one or more models responding to a specific CYP-mediated reaction, among those available in QSARINS-Chem ECO.44. The software will automatically execute the PaDEL-Descriptor software to calculate the molecular descriptors necessary to apply the chosen QSAR and calculate the predictions of interest.

The information calculated by the software (e.g., descriptor values, predictions, graphs and so on) can be exported for further analysis. It is important to highlight that QSARINS-Chem ECO.44 does not assist the user to perform “batch” predictions for multiple molecules responding to different reactions, or to generate combined analysis based on multiple predictions.

- **IVBP-Suite** uses the same models included in the QSARINS-Chem ECO.44 to calculate predictions for multiple chemicals, reactions, and ranks. In addition, it performs these tasks in batch. Moreover, the software provides results for the single QSARs as well as for combined predictions calculated from multiple models available for a specific reaction. The reactivity pathway is automatically profiled by IVBP-Suite using Toxtree with the SMARTCyp module. Results are provided as an interactive table which reports specific information on predictions and their statistics (i.e., species, assay, rank, prediction, uncertainty of the prediction, domain, and equation of the applied QSAR).