INTEGRATED COMPUTATIONAL NAM BASED ON MOLECULAR INITIATING EVENTS (MIES) AND PBK PREDICTIONS FOR EVALUATION OF THE STEATOGENIC POTENTIAL OF CHEMICALS.

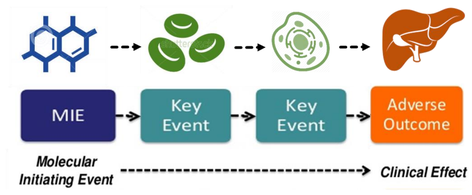
Domenico GADALETA[1]
Marina GARCÍA DE LOMANA[2]
René GECI[3]
Rafael GOZALBES[4]
Davide LUCIANI[1]
Rita ORTEGA VALLBONA[4]
Alicia PAINI[3]
Susana PROENÇA[3]
Alessandra RONCAGLIONI[1]
Eva SERRANO CANDELAS[4]
Emilio BENFENATI[1]

[1]Laboratory of Environmental Chemistry and Toxicology, Department of Environmental Health Sciences, Istituto di Ricerche Farmacologiche Mario Negri, Italy
[2]Bayer AG, Machine Learning Research, Research & Development, Pharmaceuticals, Germany
[3]esqLABS GmbH, Germany
[4]ProtoQSAR SL, Centro Europeo de Empresas Innovadoras, Parque Tecnológico de Valencia, Spain

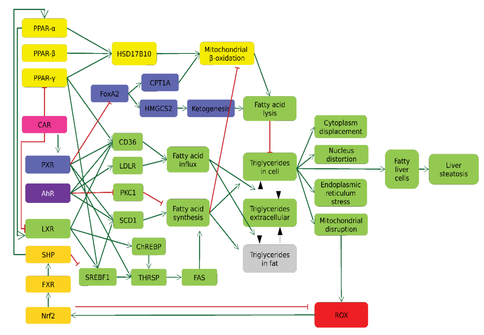
domenico.gadaleta@marionegri.it

BACKGROUND

- *Chemical-induced non-alcoholic **hepatic steatosis** is a worldwide epidemiological concern.*
- *We propose **integrated computational NAMs** to assess the steatogenic potential of chemicals, accounting for both **toxicodynamics** and **toxicokinetics.***
- *Toxicodynamics was based on the predictions of **molecular initiating events (MIE)** of the **adverse outcome pathway (AOP)** of steatosis.*
- *Toxicokinetics was estimated with an **high-throughtput (HT) PBK model.***
- *Computational approaches were developed within the EU founded project **ONTOX** and were independently applied to classify the NAM Designathon compounds (131 out of 150).*
- *The two classifications were combined to return an overall classification for the **steatogenic potential of chemicals**.*



Graphical representation of the concept of Adverse Outcome Pathway (AOP).



Schematic depiction of AOP network leading to hepatic steatosis (AOPs 57, 34, 36, 60 and 61 from the AOPWiki).

MATERIALS & METHODS

1 NAM-BASED TOXICODYNAMICS – MIE ACTIVITY QSAR

a ***QSARs to predict MIEs of steatosis** were trained on inhibition data from ChEMBL v33. Activity data were converted to binary categories (threshold IC50 =10,000 nM).*

1. ***Multi-tasking (MT) model:** simultaneous modelling of multiple endpoints to improve the final predictions.*

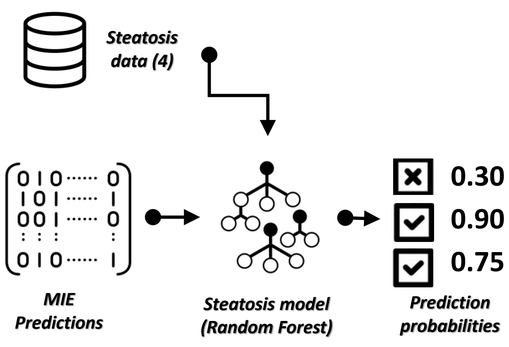
PXR | AhR | LXR | PPARα | PPARγ

2. ***Machine learning (ML) models:** individual QSAR based on ML and data transformation methods (feature selection with VSURF [2] and SMOTE [3]).*

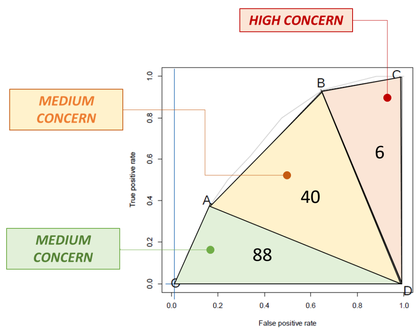
PXR* | AhR | LXR* | PPARα | PPARγ

**low performance, discarded*

b *A **random forest (RF)** model was trained on steatosis data. Predictions of the MT and ML models were used as **independent variables** to weight the contribution of each MIE to the final steatogenic outcome.*



c ***Prediction probabilities** of the RF were used to build a ROC curve.*
*The **ROC curve** was partitioned into three triangles, so that the sum of their areas maximizes the **area under the ROC (AUC).***
*Probability thresholds determined by the partition were used to classify the chemicals into **low**, **medium** and **high** risk.*



Activity classification of the EPAA compounds based on the steatosis predicting RF model

2 NAM-BASED TOXICOKINETIC – HT PBK MODEL

a ***High troughtput PBK models (HT-PBK)** were developed with the freely available, open-source modelling software **PK-Sim.**[(5)] HT-PBK was used to predict systemic availability.*

b *The HT-PBK model is parameterised with chemical input properties (**physicochemical and ADME**) predicted with various computational tools:*

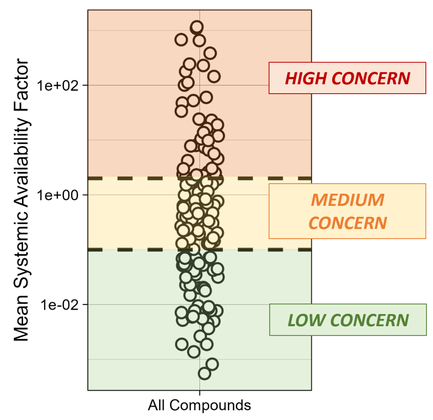
- *Water solubility*
- *Caco-2 permeability*
- *Human intrinsic hepatic clearance*
- *Fraction unbound to plasma proteins*

- *Distribution coefficient (logD) @pH =7.4*
- *Caco-2 permeability*
- *Plasma cearance*

- *Acidic and basic dissociation constant (pKa)*

- *Total cearance*

- *Distribution coefficient (logD) @pH =7.4*
- *Membrane affinity*



Bioavailability classification of the EPAA compounds based on the HT-PBK predictions

$$SAF^{*} = \frac{\text{Amount in the body (24h after last administration; after 5 years)}}{\text{Average Daily Dose}}$$

- *SAF (mean) > 200% = **high systemic availability.***
- *10% < SAF (mean) < 200% = **medium systemic availability.***
- *SAF (mean) < 10% = **low systemic availability.***

** Systemic availability factor redicted by the HT-PBK model*

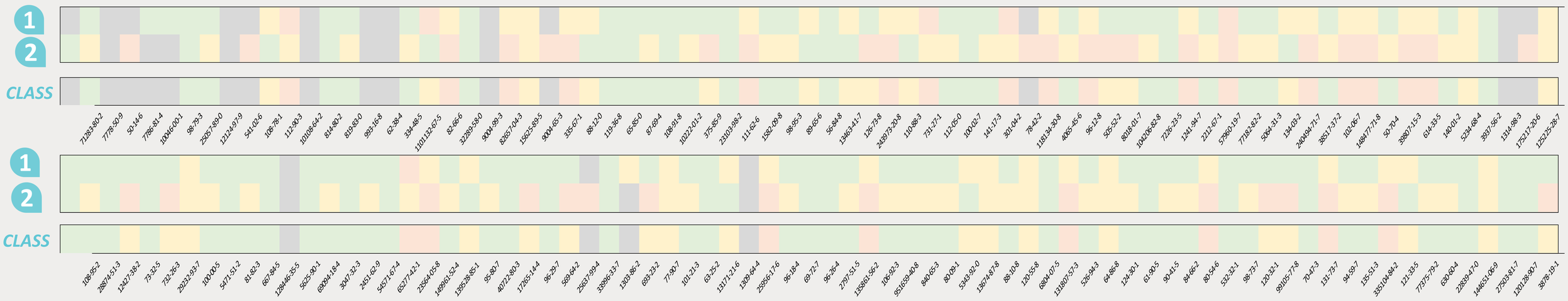
Additional details in the 'High-throughput Physiologically-based Kinetic (HT-PBK) modelling' case study report

RESULTS

*Toxicodynamics (1) and toxicokinetics (2) classifications were integrated to determine the final risk category (**CLASS**).*

- LOW CONCERN
- MODERATE CONCERN
- HIGH CONCERN
- NOT CLASSIFIED

2 Potential systemic availability (toxicokinetics)	1 Activity (toxicodynamics): High	Medium	Low
High	1	20	12
Medium	4	9	42
Low	1	10	32



References

(1) Leist, M., Ghallab, A., Graepel, R., Marchan, R., Hassan, R., Bennekou, S. H., ... & Danen, E. (2017). Adverse outcome pathways: opportunities, limitations and open questions. *Archives of toxicology*, *91*(11), 3477-3505.
(2) Genuer, R., Poggi, J. M., & Tuleau-Malot, C. (2015). VSURF: an R package for variable selection using random forests. The R Journal, 7(2), 19-33.
(3) Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, 321-357.
(4) Jain, S., Norinder, U., Escher, S. E., & Zdrazil, B. (2020). Combining in vivo data with in silico predictions for modeling hepatic steatosis by using stratified bagging and conformal prediction. Chemical Research in Toxicology, 34(2), 656-668.
(5) Willmann S, Lippert J, Sevestre M, Solodenko J, Fois F, Schmitt W (2003) PK-Sim®: a physiologically based pharmacokinetic 'whole-body' model. BIOSILICO 1(4):121–124.

This work was performed in the context of the ONTOX project (https://ontoxproject.eu/) that has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 963845. ONTOX is part of the ASPIS project cluster (https://aspiscluster.eu/).