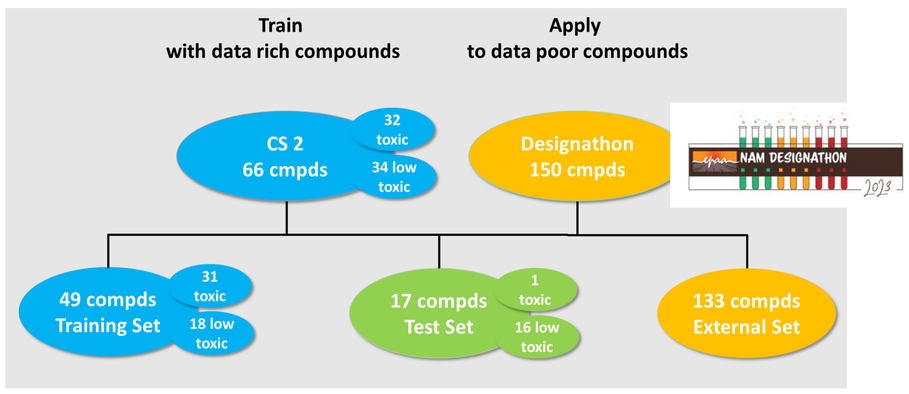
Classification of Compounds as STOT-RE based on existing high throughput NAM data

Matthias Wehr[1], Barira Islam[2], Laura I. Furlong[3], Jordi Valls-Margarit[3], Jaione Telleria-Zufiaur[3], Sylvia E. Escher[1]
[1] Fraunhofer Institute for Toxicology and Experimental Medicine; [2] Certara UK; [3] Medbioinformatics Solutions SL

Introduction

The case study evaluates the applicability of existing NAM data to classify compounds as systemically toxic for specific target organ toxicity (STOT-RE) after repeated exposure.



Approach and Ressources

66 compounds (RHT3R CS 2) are classified as **STOT-RE** based on available in vivo animal studies; subchronic oral exposure (toxic <10 mg/kg bw/d; low toxic > 1000 mg/kg bw/d). The classification approach is developed using four data sources:

- **ToxCast dataset:** AC_{50} values classified as "hits" from ToxCast database (DB) comprising about 1300 individual assays. Hits are AC50 values < cytotoxic levels. Excluded hit calls: outside of tested concentration range, data point flagged as noisy data by ToxCast. Borderline (in)active kept.
- **Chemical Effect Predictor (CEP)** uses a heterogeneous network combining chemicals, proteins, biological processes and diseases to predict drug-disease associations.
- **Physicochemical data** from PubChem and ChEMBL and predicted ADME parameters using ADMET predictor. Hepatic clearance data (***httk* DB)** available for CAS 120-55-8.
- **Kinetics:** Simcyp™ simulator was used to build **PBK models** of the compounds. Cmax and AUC predicted after single (24h)/repeated exposure (240h).

Hazard Assessment

ToxCast dataset

Trainingset 46 cmpds (30 toxic, 16 low toxic)

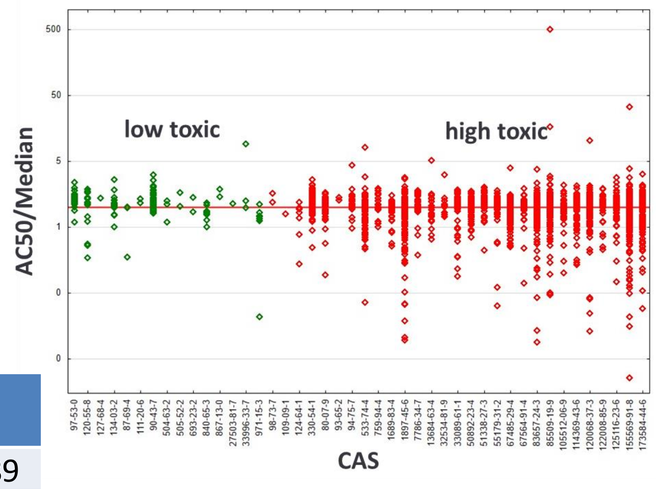
Combined ranking of compounds results in best separation of low toxic and toxic compounds:

I) Frequency – number of hit calls in all assay is higher than 5%. Data rich compounds are considered to have a higher confidence in the prediction

II) Potency - the 10th percentile of all AC_{50} values is < 5µM

RANKING		PREDICTED Toxic	PREDICTED Low toxic	
ACTUAL	High toxic	25	3	Sensitivity: 0.89
	Low toxic	1	8	Specificity: 0.88

Figure 1: *Toxic compounds show a trend toward more and sensitive hit calls*



Chemical Effect Predictor

Trainingset: 38 cmpds (28 toxic, 10 low toxic)

CEP predicts associated 7351 diseases per compound. The diseases with a predicted probability > 0.59 are considered associated diseases with the compound. A toxicity score is calculated per compounds:

$$\textbf{ToxScore} = \frac{\textbf{Proportion of pos. ass. diseases}}{\textbf{Mean probability}}$$

The 1st quartile of the *High toxic* toxicity score (1.69×10^{-3}) discriminated best low and toxic compounds.

Sensitivity 0.75
Specificity 0.70

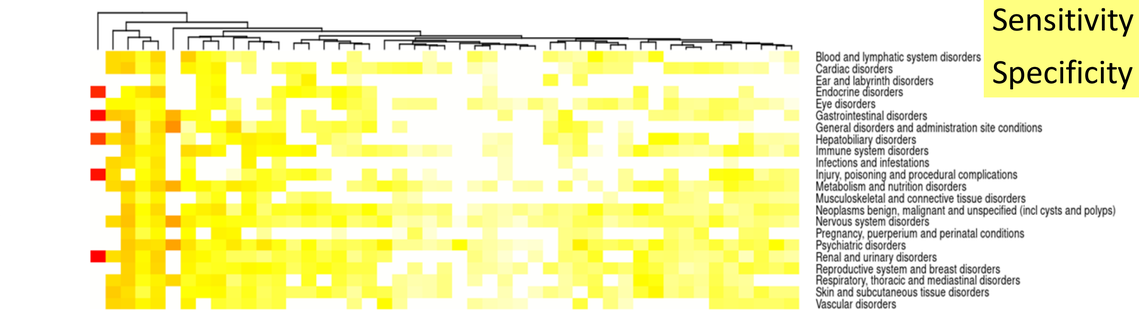


Figure 2: *CEP predicted positive association between compound and disease state (mean is shown)*

Kinetic Assessment

Internal exposure categorisation: In silico ADME data applied, except of CAS 120-55-8 to model Cmax and AUC after single or repeated exposure. All other compounds estimate **renal clearance** using a **conservative estimate** of

fu*GFR

The categorisation of these compounds might change when exp liver clearance is added to the PBK models of these compounds, as ECCS classification[1] predicts it to be cleared by liver.

CAS	Predicted Fa	Plasma Cmax (mg/L, 24 h)	Plasma Cmax (mg/L, 240 h)	AUC (24 h)	AUC (240 h)	ECCS	Internal Exposure Categorisation (single dose)
96-26-4	0.96	16.2	16.07	100.06	101.56	Class_2	Low
70-47-3	0.513	4.24	4.2	51.83	50.25	Class_4	Low
134-03-2	0.314	3.15	3.49	43.26	48.21	Class_3A	Low
819-83-0	0.67	11.7	11.14	95.03	87.41	Class_3A	Low
693-23-2	0.99	102	165	1679.6	2649.6	Class_1A	High*
505-52-2	0.91	85.9	181.91	1653.41	3480.62	Class_1A	High*
3937-56-2	0.97	9.32	17.1	117.83	299.88	Class_2	Medium*
504-63-2	0.98	1.94	1.75	96.69	106.34	Class_2	Low
120-55-8	0.99	4.4	2.07	26.06	24.94	Class_2	Low
27503-81-7	0.82	30.5	29.4	234.36	212.8	Class_1A	Medium*
33996-33-7	0.89	27.4	25.6	164.14	149.46	Class_1A	Medium*
98737	0.99	105	170.5	1711.86	2755.44	Class_1A*	High
61-90-5	0.94	18.4	18	119.17	110.75	Class_1A	Low
3047-32-3	0.99	28.6	18.6	125.99	117.07	Class_2	Low
38517-37-2	0.72	71.5	241.7	1507.93	5297.59	Class_3A	High
667-84-5	0.81	11	10.8	97.63	91.38	Class_4	Low
840-65-3	0.99	9.4	15.6	56.5	328.57	Class_2	Medium*

Resulting Classification

- The toxic compound is classified correctly by either method
- Overestimate hazard in several cases, for which hepatic clearance is likely underpredicted at the moment

Hazard (ToxCast) / Hazard (CEP)

kinetic (PBK)	Hazard (ToxCast) low	middle	high	Hazard (CEP) low	middle	high
low	134-03-2		120-55-8	61-90-5 70-47-3 96-26-4		120-55-8
middle	27503-81-7* 33996-33-7*		840-65-3*			840-65-3* 27503-81-7*
high	693-23-2* 505-52-2* 98-73-7			693-23-2* 98-73-7		505-52-2*

Green: low toxic compounds (miss classified); **Black:** low toxic compounds; **Red:** toxic compounds; ***:** might show higher hepatic clearance

Next steps

- Closer look into outliers and their mode of action
- Integrate Tier 2 data from CS2; evaluate machine learning approaches
- Measure hepatic Clint values for compounds with underpredicted clearance
- Apply derived concept to all Designathon compounds

References

1) Varma MV, Steyn SJ, Allerton C, El-Kattan AF. Predicting Clearance Mechanism in Drug Discovery: Extended Clearance Classification System (**ECCS**). Pharm Res. 2015 Dec;32(12):3785-802. doi: 10.1007/s11095-015-1749-4. Epub 2015 Jul 9. PMID: 26155985.

This project receives funding from the European Union's Horizon 2020 Research and Innovation programme under Grant Agreement No. 964537 (RISK-HUNT3R), and it is part of the ASPIS cluster.

risk-hunt3r.eu @RISKHUNT3R