In Silico Toxicity Prediction for Pre-Clinical Drug Discovery

Where are we now, and how can we make progress in the field?

Layla Hosseini-Gerami

Machine Learning Lead

3rd In Silico Toxicology Conference
29/09/22

layla@absoluteai.co.uk

What we're going to cover

Toxicity screening in the current drug discovery pipeline

In silico tools for proactive, high-volume toxicity screening

In silico tox models and industry: The big issues

In silico tox models and industry: Solutions

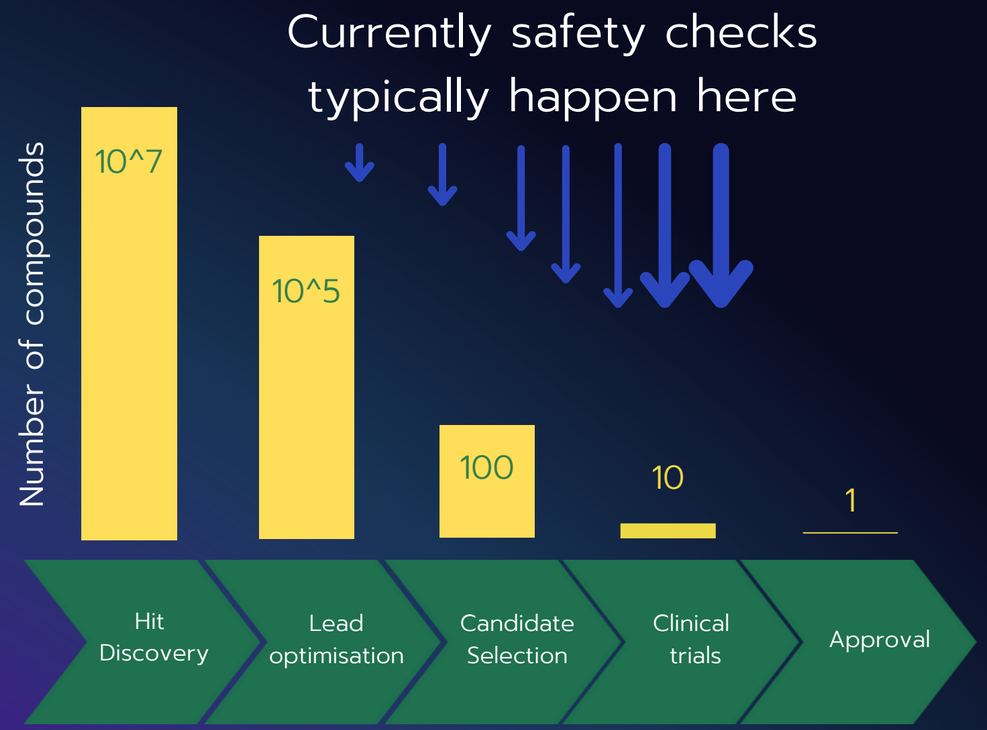
Future prospects and discussion

Why aren’t in silico tools applied more widely in industry, particularly for toxicity prediction? How can we address these issues?

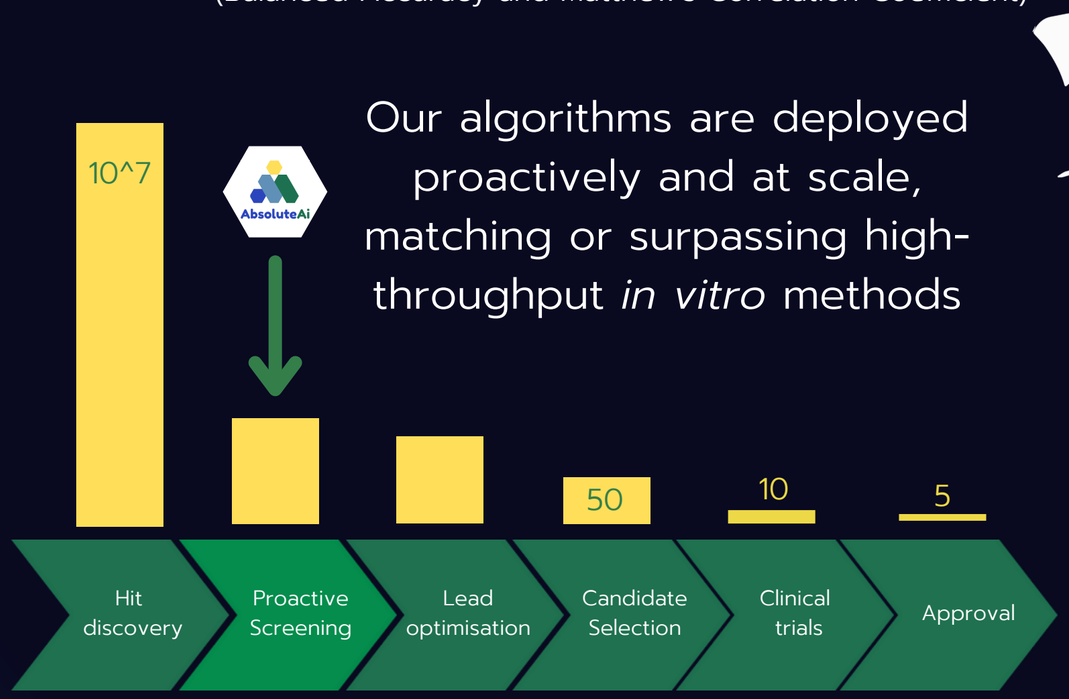
Toxicity screening: Earlier = Better

Method	BA	MCC
AbsoluteAi MitoTox	0.83	0.71
Apredica MitoMembPot [1]	0.78	0.58
Apredica MitoMass [1]	0.64	0.32

Based on a reference set of 60 compounds
(Balanced Accuracy and Matthew's Correlation Coefficient)



High failures rates in the clinic
Expensive and slow testing



Our algorithms are deployed proactively and at scale, matching or surpassing high-throughput *in vitro* methods

Fewer failures
Cheaper screening
More sustainable drug discovery

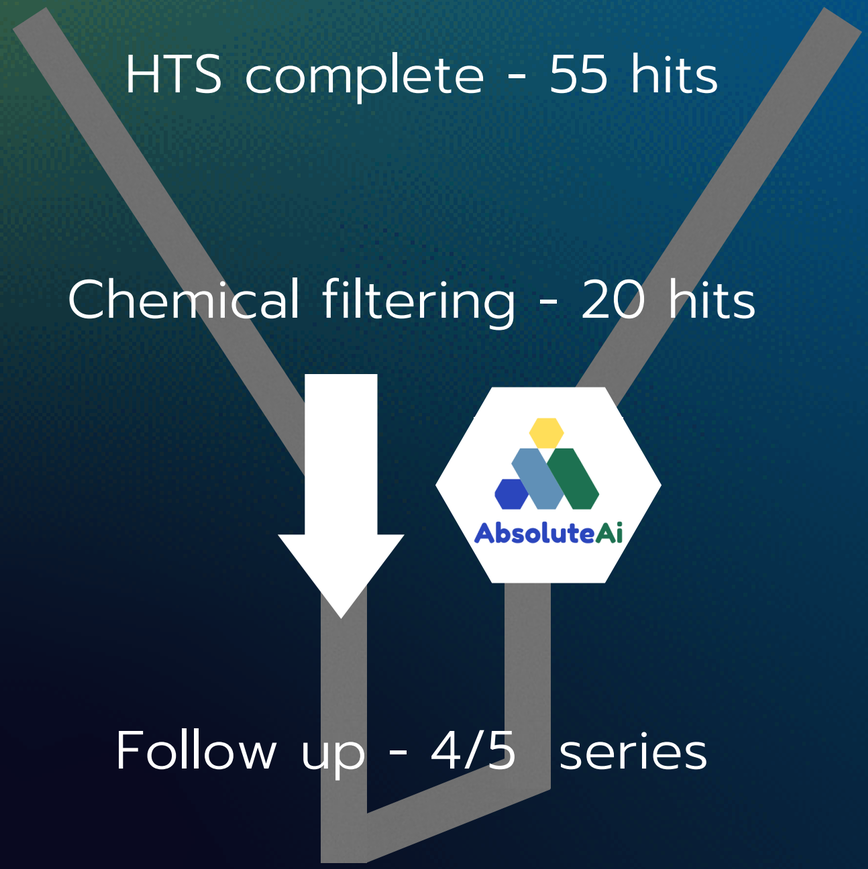
[1] Seal, S., et al. (2022) Communications Biology

Case Study: Early drug development guided by *in silico* predictions

Early biotech with a HTS output

Used traditional chemical filtering to 20 or so clusters, using 'intuition' to decide what to follow - compounds difficult to distinguish from this point

Can now use predictive toxicity as a guide away from toxic chemotypes



How can we convince industry to adopt this?

It is necessary to build models which medicinal chemists can trust, and that provide value in the decision making process i.e., not just black-box yes/no models

Problems	Solutions
• Chemist doesn't know whether the model's prediction is reliable for their compound series	• Consider model applicability domain • Calculate model confidence
• Compound may be toxic, but only at doses far higher than efficacious dose	• Build models to predict toxicity at different concentrations • Include in vitro in vivo extrapolation
• Difficult to trust a prediction coming from a black-box model - why is my compound toxic?	• Use interpretable machine learning algorithms and features • Use inherently interpretable non-ML algorithms

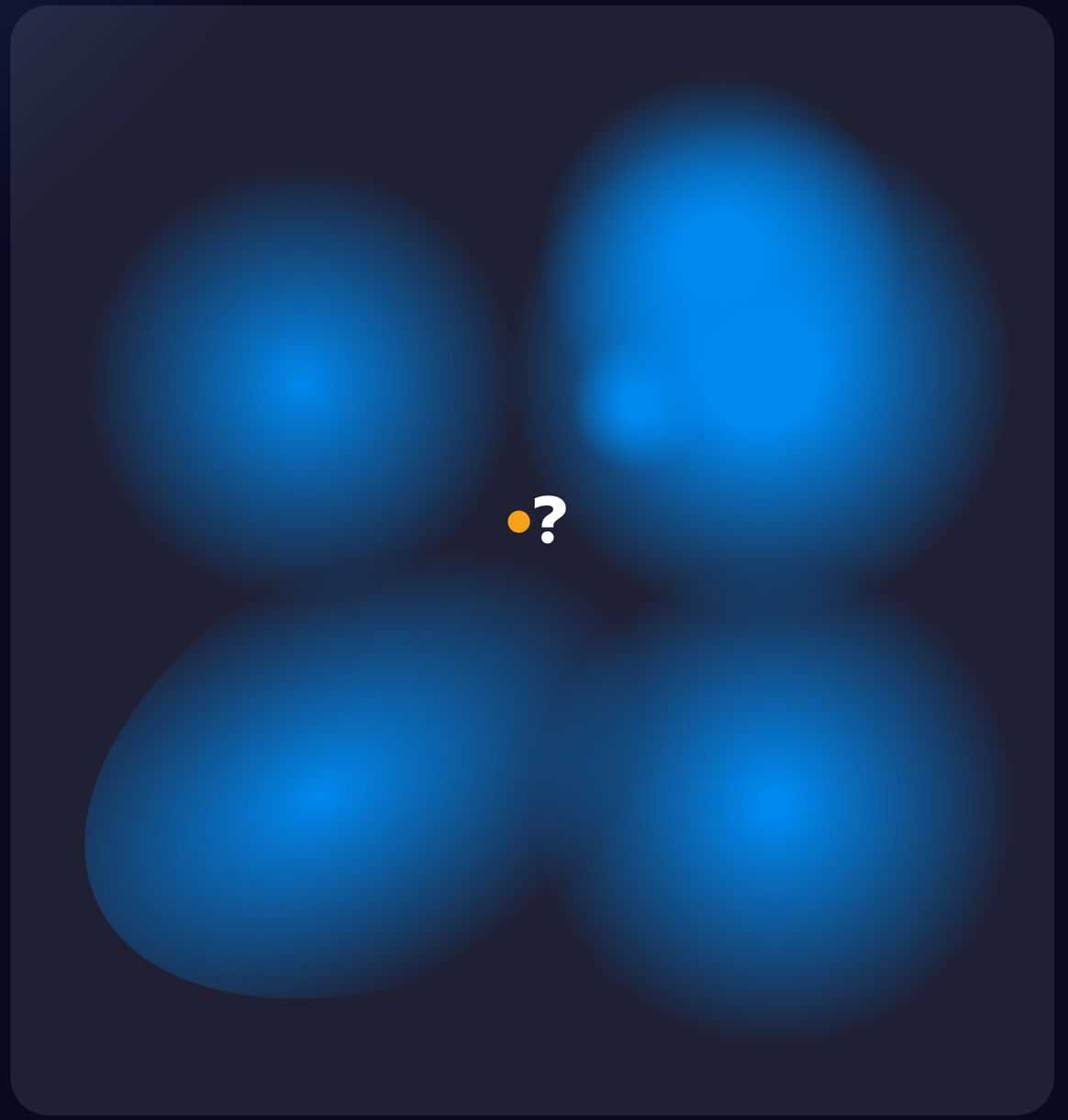
Problem 1: Chemist doesn’t know whether the model’s prediction is reliable for their compound series

Applicability domain analysis: How reliable is the prediction for my chemical series?

Applicability domain (AD): areas of chemical/biological space where model predictions are reliable

Reliability-Density Neighbourhood (RDN) [1] map of chemical space considers the local density (number of nearest neighbours) and local reliability (precision and bias) of training instances

Unseen compounds are mapped onto the RDN to assess the probability of mispredictions - can be used as a measure of trust in new predictions



Reliability-Density map of chemical space, adapted from [1]

[1] Aniceto, N., et al. (2016) Journal of Cheminformatics

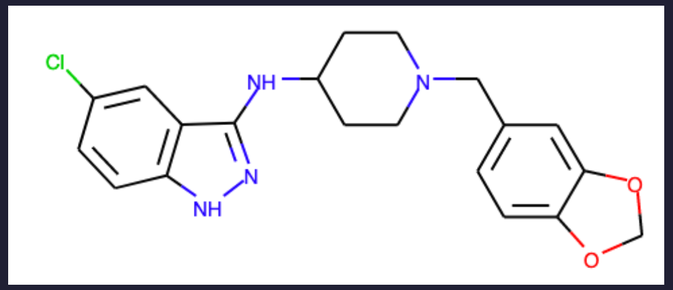
AD Analysis Example

A-777903, melanin-concentrating hormone-1 (MCHR1) antagonist

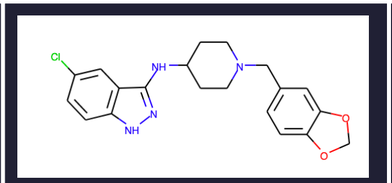
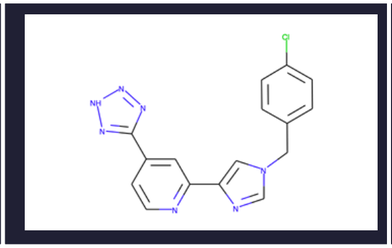
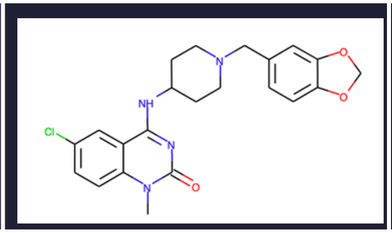
RDN and training set similarity analysis are incorporated in the PIDGINv4 [1] target prediction tool which can be used to predict on- or off-targets related to both efficacy and toxicity

Including AD analysis informs on which predictions are more or less reliable, indicated by the similarity of the nearest training set neighbours

A-777903

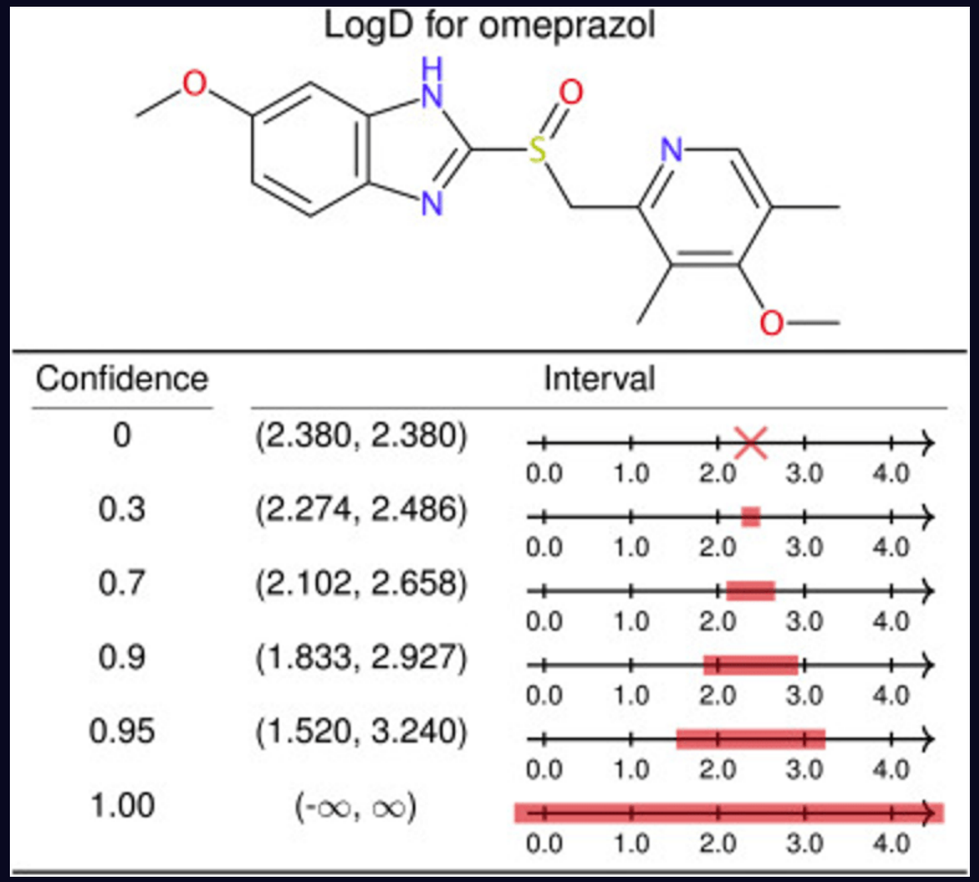


Top 3 predicted targets of A-777903 and their AD/similarity analysis results, performed using PIDGINv4

Predicted Target	Probability of Activity	RDN Score	Nearest Neighbour	Tanimoto Similarity
MCHR1	99%	99.8%		1.00
KDM4C	70%	11.5%		0.21
KCNH2/ hERG	58%	49.3%		0.63

[1] Mervin, H., et al. (2018) Bioinformatics | https://github.com/BenderGroup/PIDGINv4

Predict with confidence with conformal prediction



Conformal prediction intervals at different confidence levels for the prediction of omeprazol's LogD [1]

Conformal prediction (CP) is a framework which sits on top of traditional ML algorithms to provide mathematically valid confidence estimates for predictions [1]

CP calculates a prediction interval specific to each predicted object at a user-specified confidence level, based on the nonconformity of the predicted object

User can tune confidence levels depending on the stage of the drug discovery project - e.g., the later the predictions are applied, the higher the confidence level

[1] Alvarsson, J., et al. (2021) Journal of Pharmaceutical Sciences

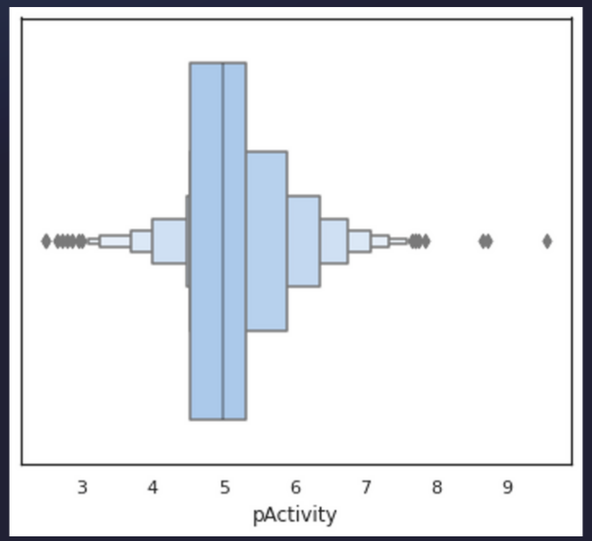
Problem 2: Compound may be toxic, but only at doses far higher than efficacious dose - these shouldn’t be thrown out at an early stage

Build models at different concentration thresholds

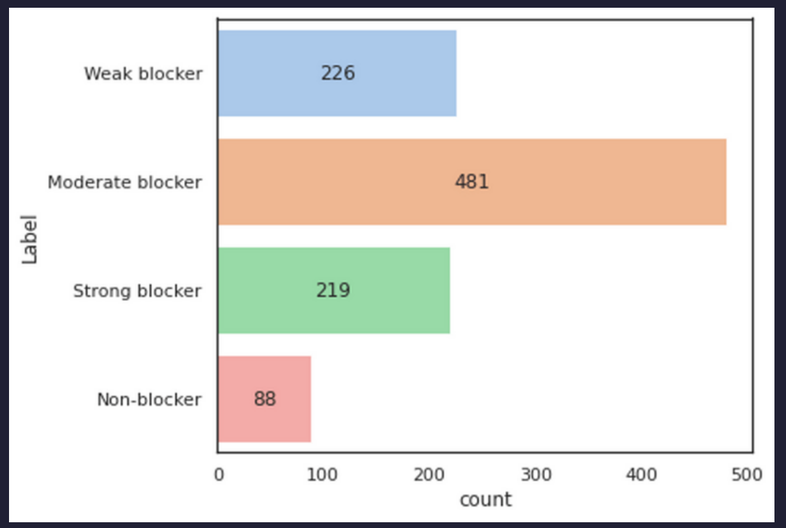
Compounds labelled as “toxic” or “non-toxic” don’t indicate likelihood of toxicity at at relevant therapeutic dose

Predictions of response at different *in vitro* concentration levels can be used to extrapolate to *in vivo* dose

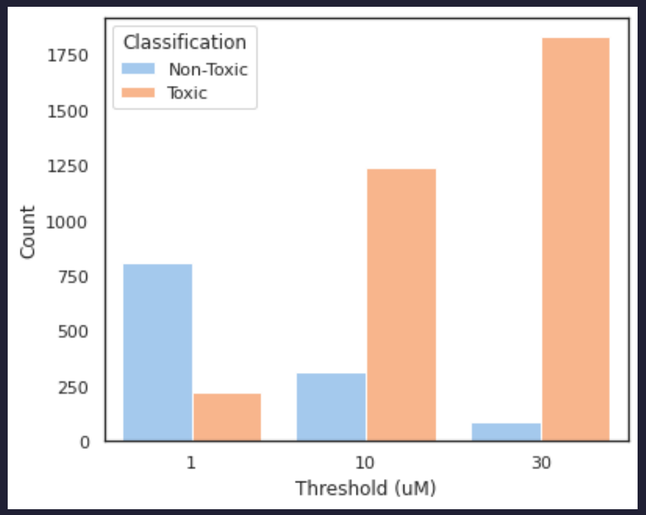
In vitro pIC50 data is represented here in terms of three different model types: regression, multi-label and binary classification at different pIC50 thresholds



Regression



Multi-label classification



Binary classification

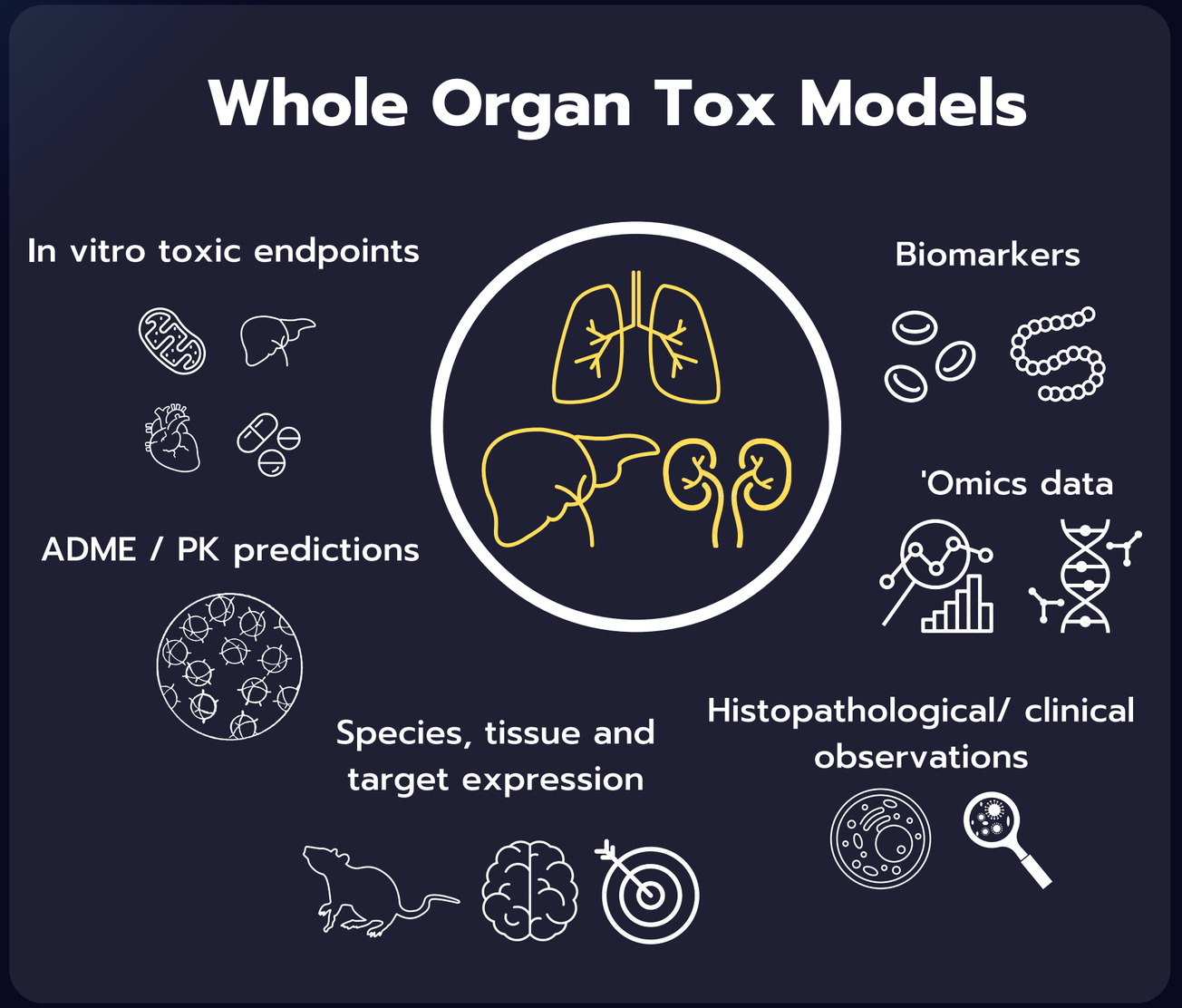
Extrapolate in vitro predictions to understand in vivo relevance

In vitro data is commonly used to build in silico tox models

In vitro-in vivo extrapolation (IVIVE) consider models' predictions in terms of in vivo relevance [1]

Exposure in AOPs (aggregate exposure pathways - AEP-AOP) will facilitate the development of more complex in silico models [2]

Other types of data can be integrated with ADME/PK predictions to build up a picture of *in vivo* organ tox



[1] Zhang, Q., et al. (2018) Frontiers in Public Health
[2] Clippinger, A., et al. (2018) Toxicology in Vitro

Problem 3: Difficult to trust a prediction coming from a black-box model.
Why is my compound toxic?

Models used in decision making must be interpretable

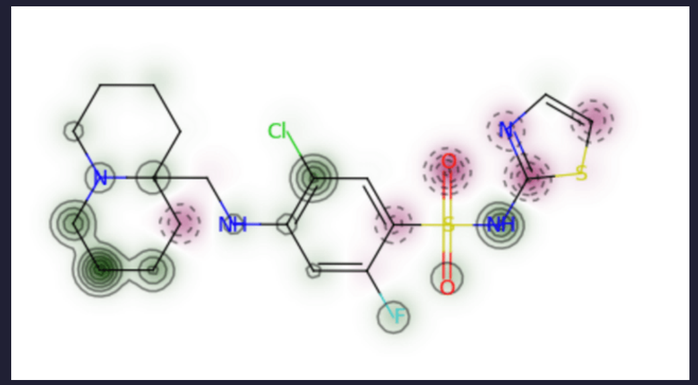
Complex deep learning architectures show good performance but are generally difficult to interpret

Tree-based methods are inherently interpretable, but more simple and thought to be less accurate

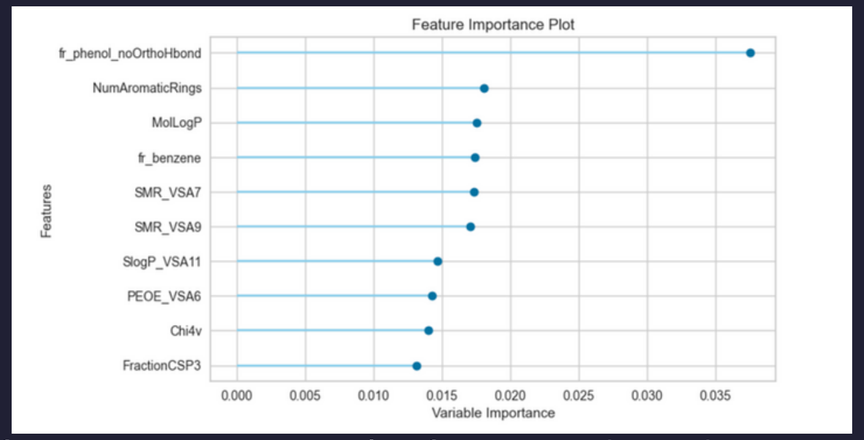
“Don’t crack a nut with a sledgehammer” - assess any performance gain when using complex, opaque models vs. simple, transparent models

Graph Convolutional Networks (GCN) treat compounds as graphs, can break down individual atomic contributions to predictions

“Meaningful” descriptors e.g., PhysChem descriptors enable interpretable feature importance analysis compared to hashed molecular fingerprints



Atomic contributions increasing (green) and decreasing (red) predicted toxicity in a GCN - visualised with RDKit



SHAP feature importance plot for a tree-based toxicity model

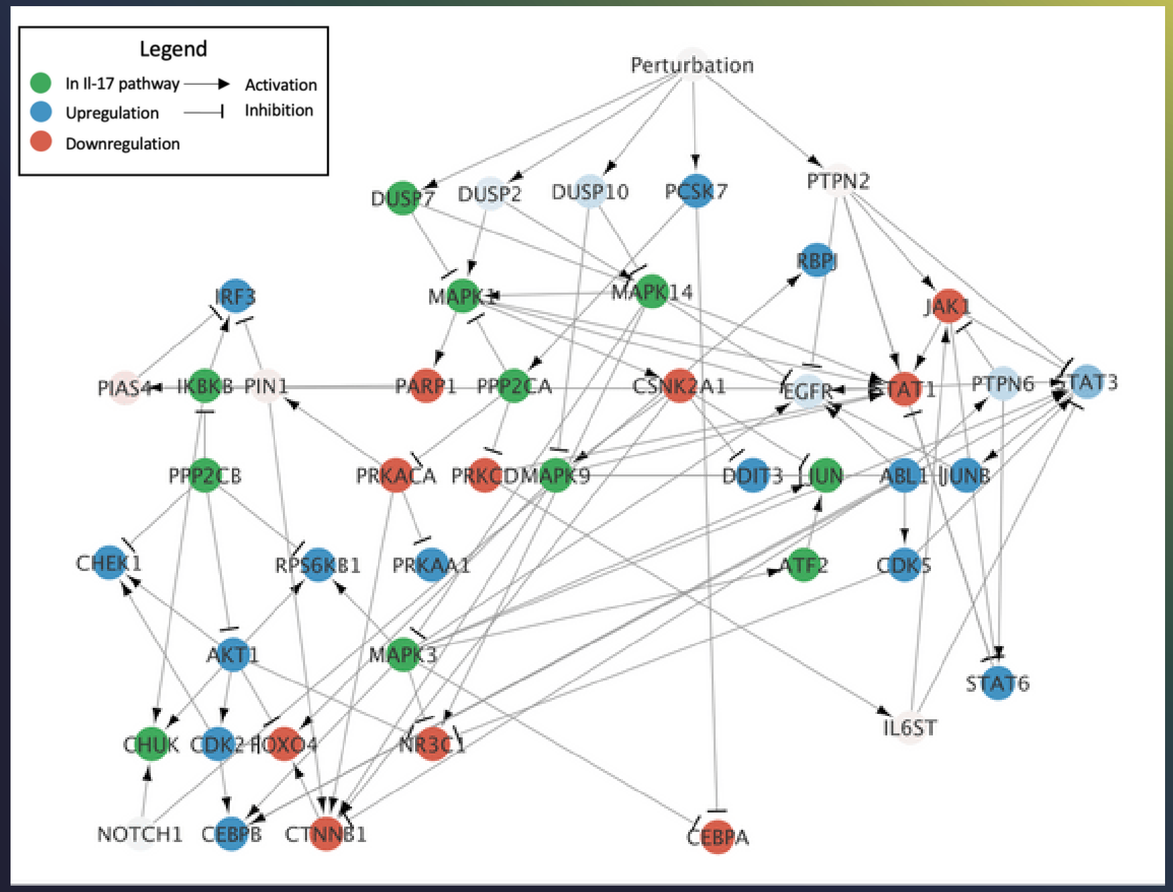
Complement ML methods with inherently interpretable non-ML methods

'Omics data can be used to investigate mechanisms of toxicity in an inherently explainable way

Causal reasoning (CARNIVAL [1] implemented in the MAVEN [2] application) applied to transcriptomics data from hepatocytes treated with high dose acetaminophen (APAP) to derive a toxicity signalling network from a hepatocyte-specific prior knowledge network

Resulting network is highly enriched for Il-17 signalling, plays a pivotal role in APAP-induced hepatotoxicity [3]

Such findings can be considered in parallel with machine learning-based predictions



Reactome Interleukin-17 Pathway: Adjusted p-value = 6.62e-6

Network of high dose APAP signalling in hepatocytes, visualised in Cytoscape

[1] Liu, A., et al. (2019) npj Systems Biology and Applications | https://github.com/saezlab/CARNIVAL
[2] Hosseini-Gerami, L., et al. (2022) BiorXiv | https://github.com/laylagerami/MAVEN
[3] Lee, HC., et al. (2018) Toxicology Letters

Tackling the big issues

How do we tackle the issues that face us, together?

Building trust and confidence with experimental scientists

Sharing data to enable better chemical space coverage of predictive models

Developments in IVIVE techniques

A call to collaborate...

layla@absoluteai.co.uk

We are looking to build on our existing partnerships and consortia across drug discovery in order to tackle the big issues that we face

Together, we can create better predictions, earlier, ultimately saving lives.

Thank you for listening

Funding for PhD work was provided by BBSRC and Eli Lilly and Company