Experiences of Predicting Acute Oral Toxicity in Industry

Richard Marchese Robinson, Surya Hari Lal, Sarah Whalley
Syngenta, Jealott's Hill International Research Centre, Bracknell, RG42 6EY, UK
richard.marchese_robinson@syngenta.com

Overview
- Knowledge of rodent acute oral toxicity (AOT) is critical for assessing the safety and registerability of plant protection product active ingredients (Al)
- Ethical imperative to reduce animal tests and extensive testing of AI leads in early stage research is not feasible
- Predictive models based on the (Quantitative) Structure-Activity Relationship (Q SAR) paradigm can support early stage research and may support future regulatory submissions

Early Stage Research
- Models can be used to
  1. perform in silico screens, guiding projects away from problematic regions of chemical space
  2. prioritise compounds for synthesis and testing
- Global models support predictions across a range of projects
- Local models are project specific – tuned to an AI lead and its derivatives

In-House Local Model Example
- Estimated posterior probability of acute toxicity at a single dose is plotted against model inputs characterising affinity for the site of toxicological action (predicted/measured K_i) and bioavailability (predicted/measured solubility)
- Compounds with low AOT probability prioritised for progression – avoiding late stage attrition
- However, building this kind of model requires a single, known mode of toxicological action

Evaluation of External Global Models
- Various Open Source and commercial software programs now provide (Q)SAR models to predict the AOT LD50 or predict a category corresponding to a range of LD50 values, e.g. GHS categories
  - Leadscope, OPERA, UL Cheminformatics Toolkit etc.
- A variety of software programs and models were identified for an evaluation of their ability to categorise compounds according to critical LD50 thresholds on a curated Syngenta in-house AOT dataset
  1. Do our early stage research compounds lie inside the applicability domains of these models?
  2. Which model is best?
  3. Is the best model good enough to support early stage research projects?

In-House Global Model?
- The option of building an in-house global model is under consideration
- This could improve the chance of our AI leads lying inside the applicability domain
- The model could be iteratively updated with new data
- Mechanistic interpretability is desired to
  - guide AI molecular design
  - support possible future regulatory acceptance
- One option – adapt a recently published framework, combining molecular descriptors and predicted protein-binding affinities for a variety of possible targets of toxicological action

Preliminary Leadscope Results
- 657 chemicals (Syngenta, PPDB) assigned GHS categories 1 – 5, or non-classified (NC), using rat AOT test data
  - Where possible, restricted to acceptable quality data (Klimisch score = 1 or 2)
  - Highly unbalanced – only 6 compounds in GHS category 1 (LD50 ≤ 5 mg/kg)
- Multi-class predictions made using consensus of Partial Logistic Regression and structural alert models
  - The most conservative (lowest GHS category) prediction is returned
  - The results are encouraging
    - Only 21 inconclusive predictions had to be discarded
  - The balanced accuracy (45%) is much higher than would be expected due to chance (17%)
- On average, 90% of chemicals in a given category are assigned to the correct or a more conservative class – the assignment required in a regulatory setting

Possible Workflow to Support Regulatory Submissions

References
3. Bercu et al. A cross-industry collaboration to assess if acute oral toxicity (Q)SAR models are fit-for-purpose for GHS classification and labelling. (Manuscript in Preparation)

Classification: PUBLIC