# Integrating Adverse Outcome Pathways (AOPs) and High Throughput in vitro Assays for Better Risk Evaluations, a Study with Drug Induced Liver Injury (DILI)

Kapil Khadka, Minjun Chen, Zhichao Liu, Weida Tong, Dong Wang

*Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, FDA*

Dong Wang
Dong.wang@fda.hhs.gov

## Background and Approaches

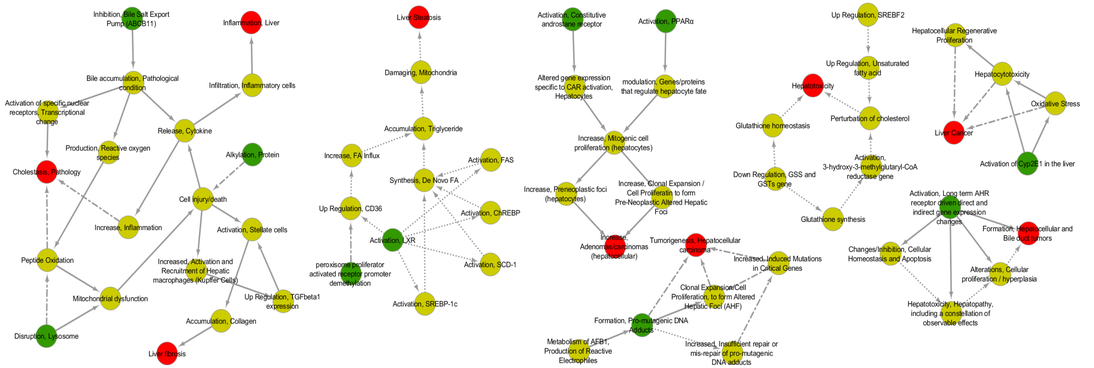
***Background***

- High throughput in vitro assays like those developed for ToxCast/Tox21 and L1000 provide a valuable source of data for a wide range of biological functions
- Direct use of all high throughput endpoints is not effective, especially with a small sample size
- The adverse outcome pathway (AOP) framework provides a convenient way to encode and communicate mechanistic knowledge
- We aim to integrate mechanistic knowledge with multiple *in vitro* assay data sources to develop predictive models for the risk of drug induced liver injury (DILI)

**Approaches**

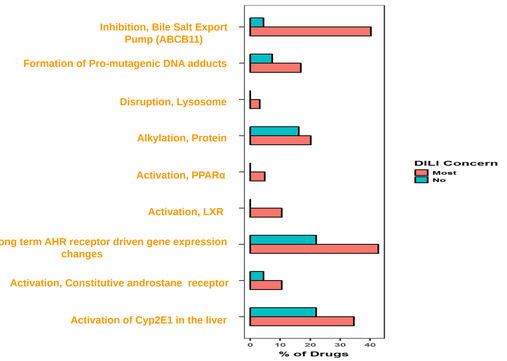
- Map drugs with known DILI properties to AOPs using information in DrugBank
- Build high-dimensioial predictive models for DILI risk with predictors selected with the aid of AOP networks

## AOPs with hits by most-DILI-concern drugs




- AOPs with hits by most-DILI-concern drugs form seven distinctive AOP networks
- The red nodes are adverse outcomes (AOs), the green nodes are molecular initiating events (MIEs), and the yellow nodes are key events (KEs)

## Frequency of hits by Most-DILI-concern and no-DILI-concern drugs




Liver toxicity related AOPs have varying frequencies of hits by most-DILI- concern and no-DILI-concern drugs

## Build predictive models with high throughput assays and AOP networks

***Selection of predictors aided by AOP networks***

- Liver toxicity related AOPs in AOPwiki and literature were used to identify 64 potential predictors, including endpoints for nuclear receptor binding, gene expression and cellular function
- These were matched to 12 assays in Tox21 and 31 gene expression measures from L1000
- Three drug properties for daily dose, logP, and reactive metabolite (RM) formation are also included for a total of 46 predictors

**Logistic regression model with elastic net penalty**

- The response (*y*) indicates either most-DILI-concern (92 drugs) or no-DILI concern (64 drugs). The vector *x* encode the 56 predictors

$$log\frac{P(y=1|X=x)}{P(y=0|X=x)} = \beta_0 + \beta^T x$$

- The penalty parameter λ is selected by cross-validation

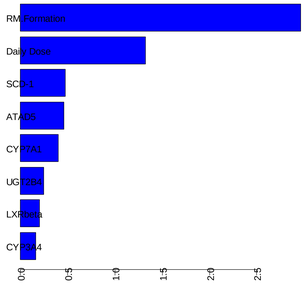
$$(\hat{\beta}_0, \hat{\beta}) = \min_{(\beta_0,\beta)} -\frac{1}{N}\sum_{i=1}^{N}\left[y_i(\beta_0 + x_i^T\beta) - \log\left(1 + e^{\beta_0 + x_i^T\beta}\right)\right] + \lambda\left[\frac{(1-\alpha)\|\beta\|_2^2}{2} + \alpha\|\beta\|_1\right].$$

**Disclaimer**
**The information in these materials is not a formal dissemination of the U.S. Food and Drug Administration.**

## Performance of the penalized logistic regression aided by AOP networks

| Model Fit | | Actual | |
|---|---|---|---|
| | | No-DILI-Concern | Most-DILI-Concern |
| Fitted | No-DILI-Concern | 45 | 4 |
| | Most-DILI-Concern | 9 | 88 |

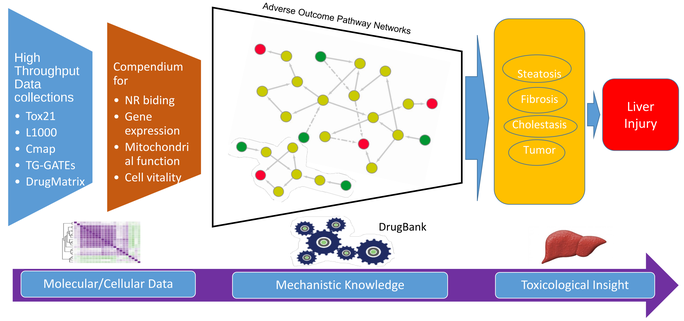
| Cross-validation | | Actual | |
|---|---|---|---|
| | | No-DILI-Concern | Most-DILI-Concern |
| Predicted | No-DILI-Concern | 39 | 11 |
| | Most-DILI-Concern | 15 | 81 |




Predictors with the largest coefficients in the penalized logistic regression model

- On model fit, the penalized regression model obtained a sensitivity of 0.96, a specificity of 0.83, with the accuracy being 0.91.
- Under cross-validation, it obtained a sensitivity of 0.88 and a specificity of 0.72, with the accuracy being 0.82.
- In comparison, the model fit accuracy using only three drug properties (daily dose, logP, and RM formation) is 0.84, that using only in vitro assays is 0.63. Both are lower than that of the penalized logistic regression model using all selected predictors (0.91).

## Conclusions




- Mechanistic knowledge is of critical importance even with the availability of high throughput in vitro data
- Our results confirmed the utility of AOPs as a easily accessible source of mechanistic knowledge
- AOPs can be effectively utilized to select endpoints from in vitro assays as a dimension-reduction tool
- Integrating multiple data sources with AOPs is promising
- The linear score for DILI potential has been shown to be significantly associated with higher reporting rates in the FAERS data