Modeling the Risk of Drug-Induced Liver Injury with Adverse Outcome Pathways and Bayesian Networks

Thilini Mahanama, Arpan Biswas, Dong Wang

Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, 3900 NCTR Rd, Jefferson, AR 72079

Introduction

The development of high throughput in vitro assays has a profound impact on toxicological assessment with the potential to lead to more efficient, accurate, less time-consuming, and less animal-intensive testing. However, utilizing numerous in vitro tests in actual risk assessment processes is still a significant challenge. Since the adverse outcome pathway (AOP) framework, as rich source of mechanistic knowledge, is a promising tool to select and structure in vitro assays in predictive models for toxicity, we utilize knowledge encoded in AOPs to build a predictive model for drug-induced liver injury (DILI) using a Bayesian network approach. We show that by incorporating mechanistic knowledge in our model, we can achieve competitive performance for predicting DILI risk with only a small number of in vitro assays in comparison with other machine learning based methods.

Data

The DILI risk classification was obtained from Liver Toxicity Knowledge Base. Nuclear receptor binding assay results were obtained from Tox21 project and gene expression assays were from L1000 project, both were processed as z-scores. AOPs were extracted from AOPwiki.

Identifying the AOPs related to DILI

We focused on three types of liver injuries: fibrosis, cholestasis, and steatosis. For each type of injury, relevant AOPs were extracted from AOPwiki (https://aopwiki.org/). As some AOPs share common elements, these AOPs form networks for each type of injury. We further augmented the networks with results in the literature. These AOP networks represent available mechanistic knowledge for molecular events that lead to liver injury by drugs and other chemicals.

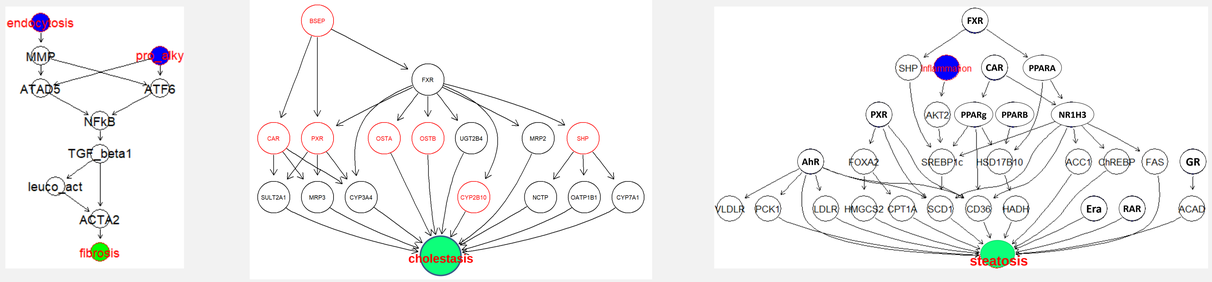

Figure 1: Updated AOPs for fibrosis, cholestasis, and steatosis (left to right) based on AOP-wiki and literature. For simplicity, we focus on molecular events in the AOP networks and omitted other nodes not relevant to the model.

We further constructed a directed acyclic graph (DAG) for modeling DILI with a Bayesian network model. Based on the theory of Bayesian networks, we only retained the nodes (molecular assays) that forms the Markov blanket of the nodes representing the three types of injury (fibrosis, cholestasis, steatosis). Thus, the number of nodes for molecular assays has been significantly reduced in this step.

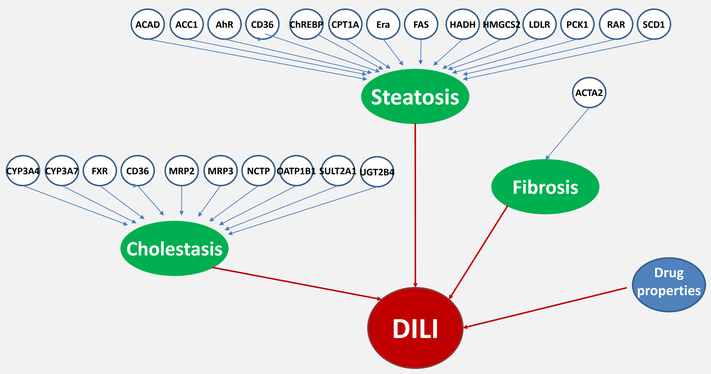

Figure 2: The DAG corresponding to the Bayesian network model for DILI. Only molecular assays that form the Markov Blanket of the three types of injuries were retained in the DAG.

Predicting the Risk of DILI based on AOPs

We constructed two Bayesian network models based on the assays in Figure 2 while using L1000 and Tox21 data for gene expression and nuclear receptor binding. We also incorporated three drug properties (daily dose (*DD)*, lipophilicity or *logP*, and reactive metabolite formation (*RM*)) in the model. We evaluated the performance of models in predicting the risk of DILI for well curated drugs.

Model 1: Bayesian model for DILI based on Fibrosis (FI), Cholestasis (CH), Steatosis (ST), and drug properties. In this model, the absolute z-score for the in vitro assays relevant to each injury type is averaged to generate a single score for each type (FI, CH, and ST). These three scores and the drug properties form the covariate *X* in a probit model for DILI risk:

$$\mathbf{P}(Y = \mathbf{1}|\boldsymbol{X}) = \Phi(-\mathbf{2.99} + \mathbf{0.17}\ \mathrm{FI} + \mathbf{0.49}\ \mathrm{CH} + \mathbf{0.21}\ \mathrm{ST} + \mathbf{0.99}\ \boldsymbol{DD} + \mathbf{0.04}\ \boldsymbol{logP} + \mathbf{1.81}\ \boldsymbol{RM})$$

where *Y* is the indicator for DILI risk and Φ is the cumulative distribution function for the standard normal distribution. We used uninformative priors for all coefficients when fitting the model. Markov Chain Monte Carlo simulations were carried out with Rjags package.

Model 2: Bayesian model directly based on all in vitro assay scores and drug properties. In this model, the molecular assay z-scores are directly included in the probit model together with three drug properties. A shrinkage prior is imposed on regression coefficients for molecular assay z-scores for regularization purposes. Markov Chain Monte Carlo simulations were carried out with Rjags package.

Model Performance:

Table 1: Confusion matrices for Model 1 and Model 2

Model 1		Predicted	
		No-DILI-Concern	Most-DILI-Concern
Actual	No-DILI-Concern	49	15
	Most-DILI-Concern	5	77

Model 2		Predicted	
		No-DILI-Concern	Most-DILI-Concern
Actual	No-DILI-Concern	44	12
	Most-DILI-Concern	10	80

Figure 3: Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves for Model 1 and Model 2

Model Comparison:

	Model 1	Model 2
Accuracy	0.8630	0.8493
Sensitivity	0.8370	0.8696
Specificity	0.9074	0.8148
Area under ROC	0.928	0.925
Area under PR	0.947	0.959

Conclusion

- Our Bayesian network model using AOPs to inform model structure achieved modeling accuracy competitive with published models in the prediction of DILI risk.
- By incorporating mechanistic knowledge, our model is far more parsimonious (requires far less in vitro assays) than other machine learning methods and is thus easier to implement in the practice.

Disclaimer

The opinion expressed in this poster is of the authors. It does not necessarily represent the position of the U.S. Food and Drug Administration.