A Bayesian multi-layered model to predict mechanisms, types, and severity of drug-induced liver injury.

Elizaveta Semenova[1], Dominic Williams[2], Stanley E Lazic[1]

Data Science and Quantitative Biology group, AstraZeneca, Cambridge UK[1] Clinical Pharmacology and Safety Sciences, AstraZeneca, Cambridge UK[2]

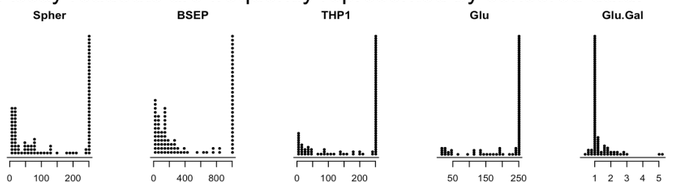
Background

Drug-induced liver injury (DILI) is a major cause of attrition in drug development and a common reason for withdrawing a drug from the market [1]. Predicting clinical liver toxicity is difficult due to its multi-mechanistic nature and chemical properties of the drug. Retrospective analysis has shown that preclinical animal studies fail to make correct predictions in about 45% of clinical trials [2]. Most current in vivo and in silico methods only predict the presence or absence of DILI, and not the severity. In addition, the mechanisms and types of injury are not used to make predictions. We therefore developed a multi-layered Bayesian model where we use preclinical in vitro assay results and physical/chemical properties of the drugs to predict the mechanisms of toxicity, use the mechanisms to predict the type of liver injury, and then combine the type of injury with the clinical dose of a drug to predict the severity of injury. For new test compounds, the model can therefore predict the mechanism and type of injury, in addition to the severity.

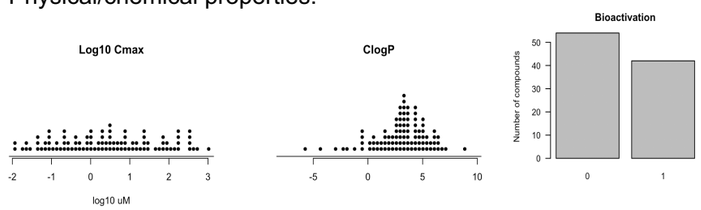
Data

Predictors

Assay readouts are frequently represented by censored data:

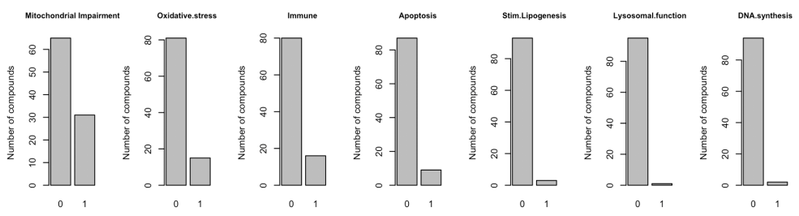


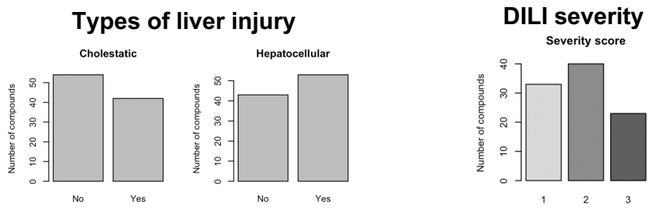
Physical/chemical properties:



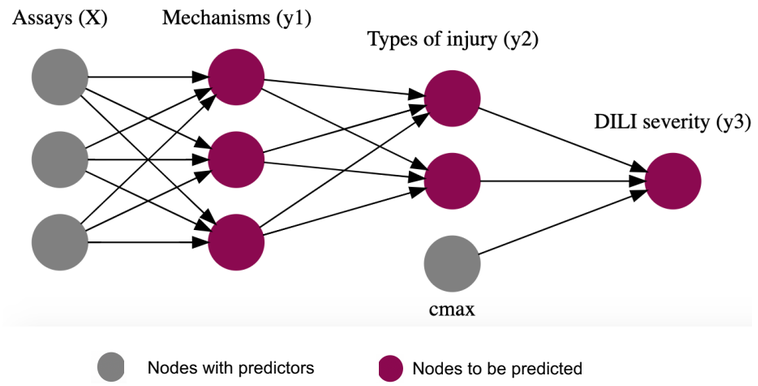
Mechanisms

Data for most of the mechanisms is highly imbalanced. Therefore we kept two nodes in the layer of mechanisms: mitochondrial impairment and other.





Multi-layered model structure



Model formulation

Likelihood

DILI severity was modelled in Stan [3] via the ordered logistic regression with three classes. Mechanisms and types of injury were modelled via logistic regression. Interactions between predictors were included on all levels of the model.

$$y_1 \sim \text{BernoulliLogit}(\eta_1)$$
$$y_2 \sim \text{BernoulliLogit}(\eta_2)$$
$$y_3 \sim \text{OrderedLogistic}(\eta_3, c)$$

Latent linear predictors

$$\eta_1 = X\beta_1 + b_1$$
$$\eta_2 = \tilde{\eta}_1\beta_2 + b_2$$
$$\eta_3 = (\tilde{\eta}_2, c_{\max})\beta_3$$

The operation $\tilde{\eta}$ expands the matrix of predictors η into the matrix of predictors and their pairwise products to account for interactions; $(\tilde{\eta}_2, c_{\max})$ denotes horizontal stacking of the matrix $\tilde{\eta}_2$ and the vector $c_{\max}$.

Priors

Horseshoe shrinkage prior [4] was applied to the effects of predictors on the mechanisms of liver injury β_1. Normal priors were given to β_2 and β_3.

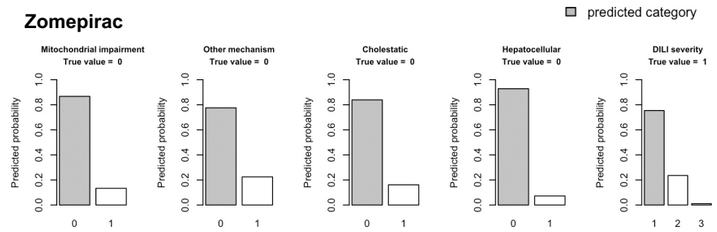
Results

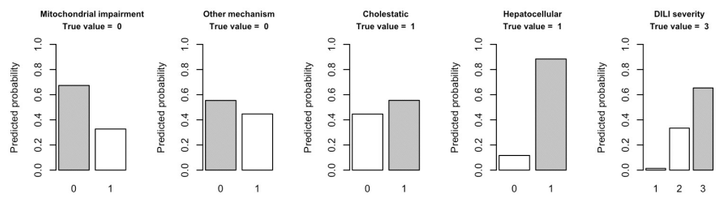
Table 1 Balanced accuracy across all predicted nodes:
By the standards of the field, the model achieves good accuracy.

Node	Balanced accuracy
Mitochondrial impairment ($y_{1,.}$)	0.70
Other mechanisms ($y_{1,.}$)	0.61
Cholestatic type of injury ($y_{2,.}$)	0.75
Hepatocellular type of injury ($y_{2,.}$)	0.90
DILI severity (y_3)	0.95*

(*) Computed as balanced accuracy for binary data since it is more important to distinguish between safe/unsafe compounds, rather than between the two unsafe categories.

Predictions





References

1. Drug-induced liver injury: recent advances in diagnosis and risk assessment. Kullak-Ublick, Gerd A., et al. Gut 66.6 (2017): 1154-1164.
2. Concordance of the toxicity of pharmaceuticals in humans and in animals. Olson, Harry, et al. Regulatory Toxicology and Pharmacology 32.1 (2000): 56-67.
3. Stan: A probabilistic programming language. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A (2017). Journal of Statistical Software 76(1). DOI 10.18637/jss.v076.i01
4. Sparsity information and regularization in the horseshoe and other shrinkage priors. Piironen J, Vehtari A. Electronic Journal of Statistics 11.2 (2017): 5018-5051.

Acknowledgements
We are thankful to A. Gelman and A. Vehtari for useful discussions of our model.