Functional interrogation of US EPA's Toxcast database for identification of hepatotoxic drugs

Sricharan Bandhakavi, Research Fellow, EBTC

Minjun Chen, US FDA NCTR

Katya Tsaioun, Evidence-based Toxicology Collaboration at Johns Hopkins Bloomberg School of Public Health

About EBTC

What is EBTC?

EBTC is an international collaboration of science, regulatory and industry leaders formed to establish and coordinate evidence-based, transparent toxicology and safety assessment methods to improve the risk assessment standards for regulatory decision making.

EBTC's Vision:

Evidence-based toxicology is the standard used to ensure public health, a healthy environment and a sustainable future.

EBTC's Mission:

Bring together the international toxicology community to facilitate use of evidence-based toxicology to inform regulatory, environmental and public health decisions.

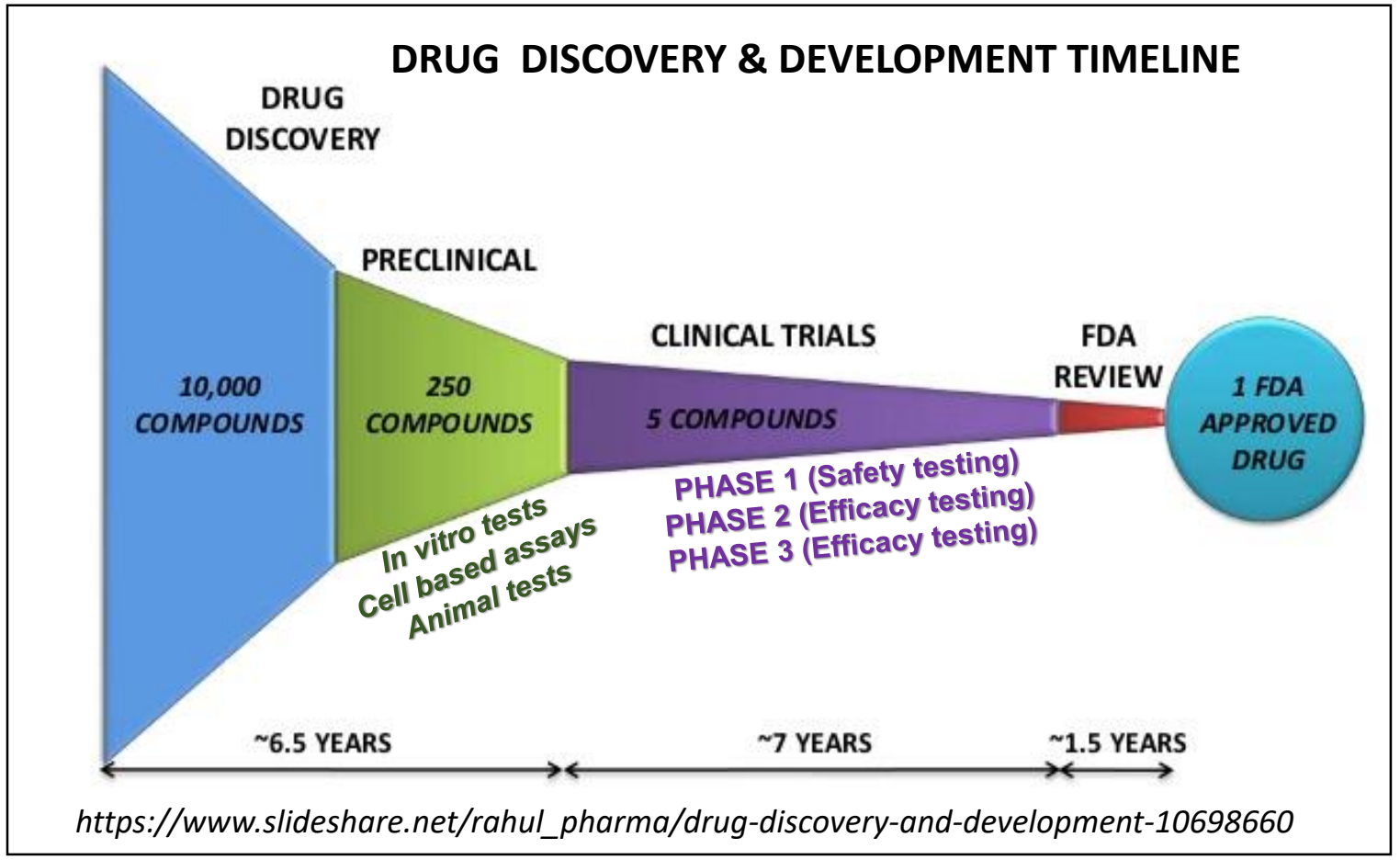
EBTC Funding:

Anonymous Charitable Foundation (87%), Beagle Freedom Prize (10%), ExxonMobil Foundation (3%)

www.EBTox.org

ebtc
Evidence-based Toxicology Collaboration

How to eliminate putatively liver toxic drugs from going into clinical trials/market ?



https://www.slideshare.net/rahul_pharma/drug-discovery-and-development-10698660

Test

Leverage preclinical/Toxcast data for identifying drugs with ***potential toxicities to liver***

Why Toxcast?

Toxcast* assays represent > 1100 endpoints that cover > 200 high-level cell responses (gene targets/pathways)

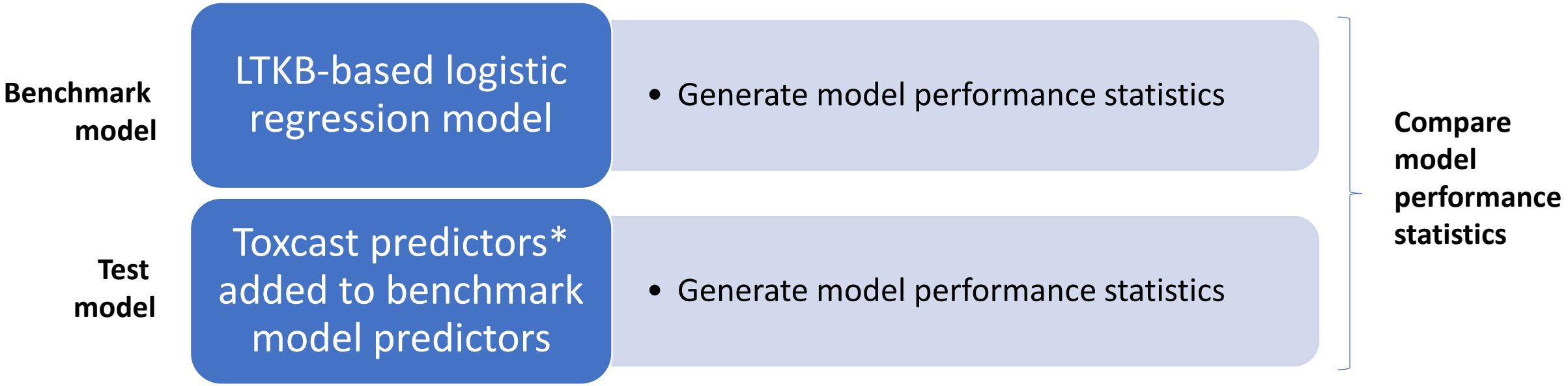
Toxcast has in vitro/cellular assay data for ~ 2000 chemicals including > 500 FDA approved pharmaceutical drugs

Preliminary data indicates diagnostic potential of Toxcast assay targets/pathways for identification of “mostDILI” drugs (data not shown)

*https://www.epa.gov/chemical-research/toxicity-forecasting

Study aim

- Determine what value, if any, does Toxcast provide to existing models of DILI *(prediction of "mostDILI" vs "other" DILI class drugs)*



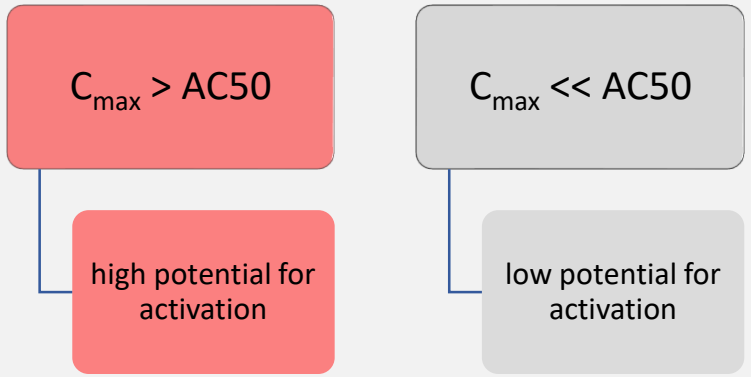
Identified from TOXCAST db2

Toxcast data retrieval/preparation

Obtain AC50 values from <= 1193 Toxcast^{db2} assays for 503 drugs	• Data containing AC50 values loaded into R
Merge Toxcast data with C_{max} & DILI information from LTKB*	• Based on 350 drug matches with C_{max}, 350RX1197C data frame created. Merged also DILI class labels.
Impute missing AC50 values for drug/assay combinations	• 79% of data points set to 10^6 (indicates no activation of assay by drug)
Calculate *Normalized Activation Score (NAS)*** for each drug/assay combination	• Select all drugs with medium NAS value or higher for further modeling

FDA's Liver Toxicity Knowledge Base used to extract C_{max} valuessee next slide for description of NAS*

Normalized Activation Score for stratification of ToxCast test/targets' activation potential in humans



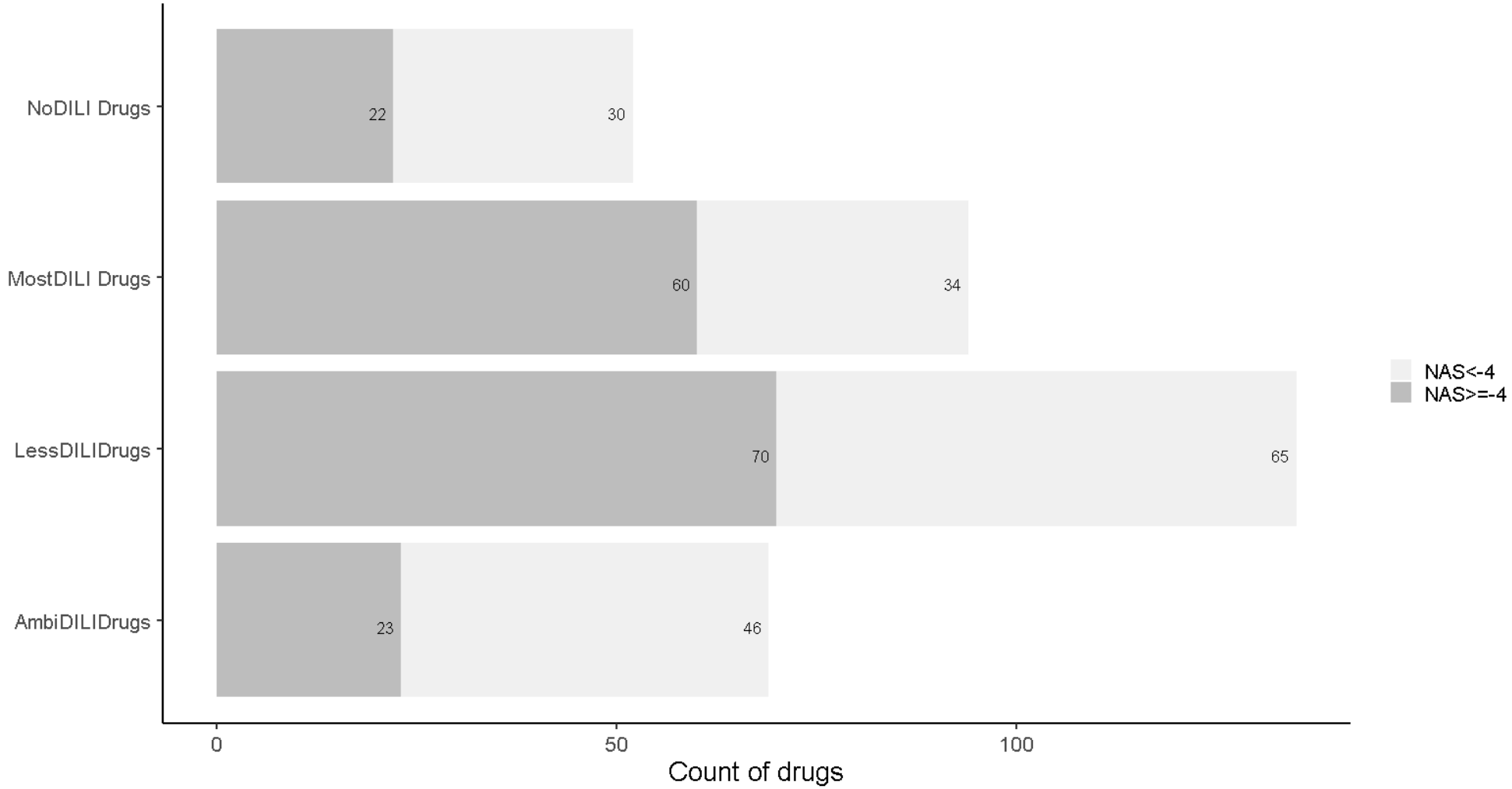
$$(Normalized)\ Activation\ Score$$

$$NAS = \frac{C_{max} - AC50}{C_{max}}$$

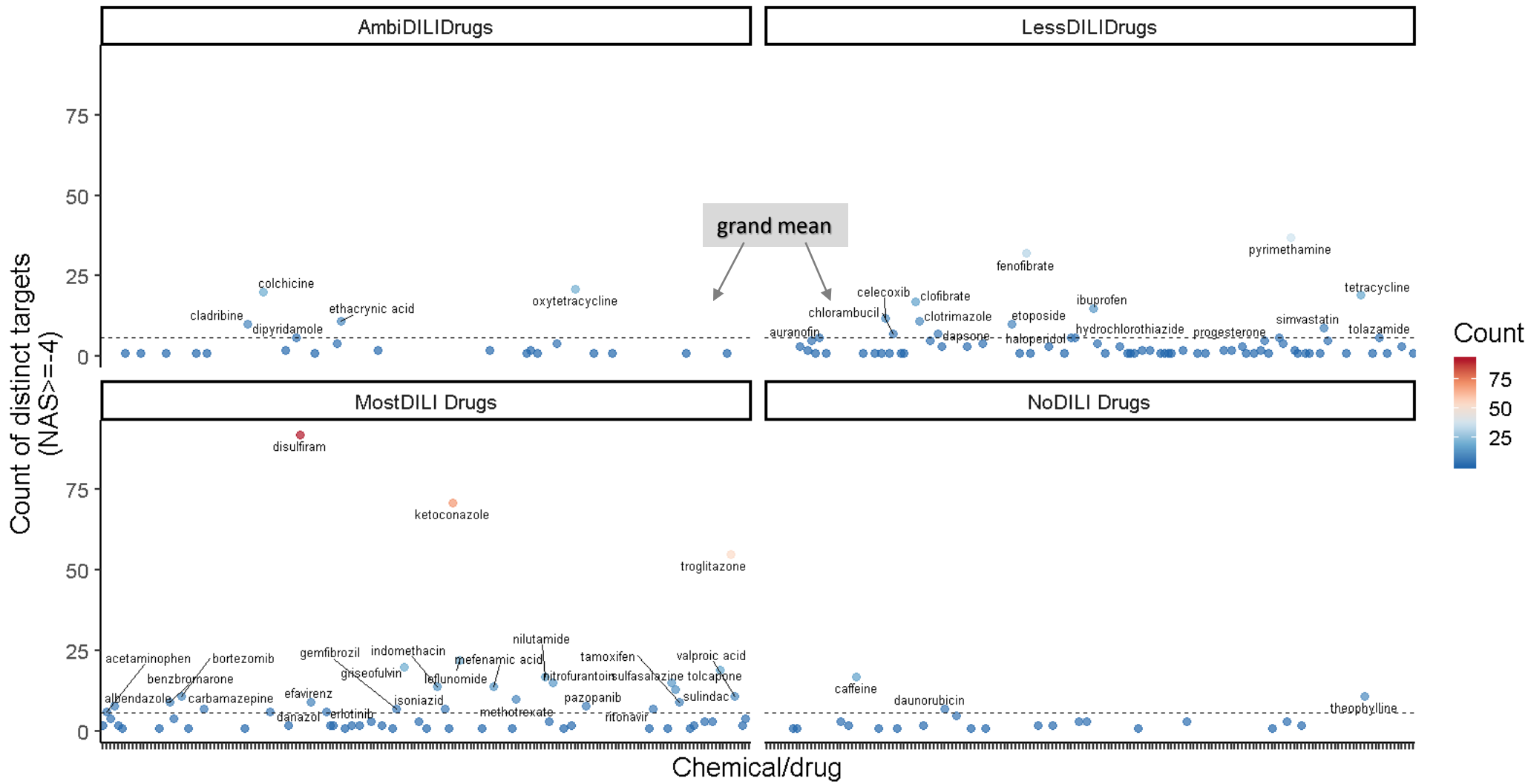
AC50 = (drug) concentration at 50% max response in ToxCast assay collection

Normalized activation score (NAS)	Indication
NAS>=0	C_{max} > AC50; highest activation potential
0>NAS>-4	C_{max}< AC50; modest activation potential
NAS<-4	C_{max}<<< AC50; lowest activation potential

175/350 drugs selected based on medium NAS value of >=-4 for further analysis/modeling



Pharmacological activity of drugs by DILI class shows relatively low # of Toxcast targets/drug

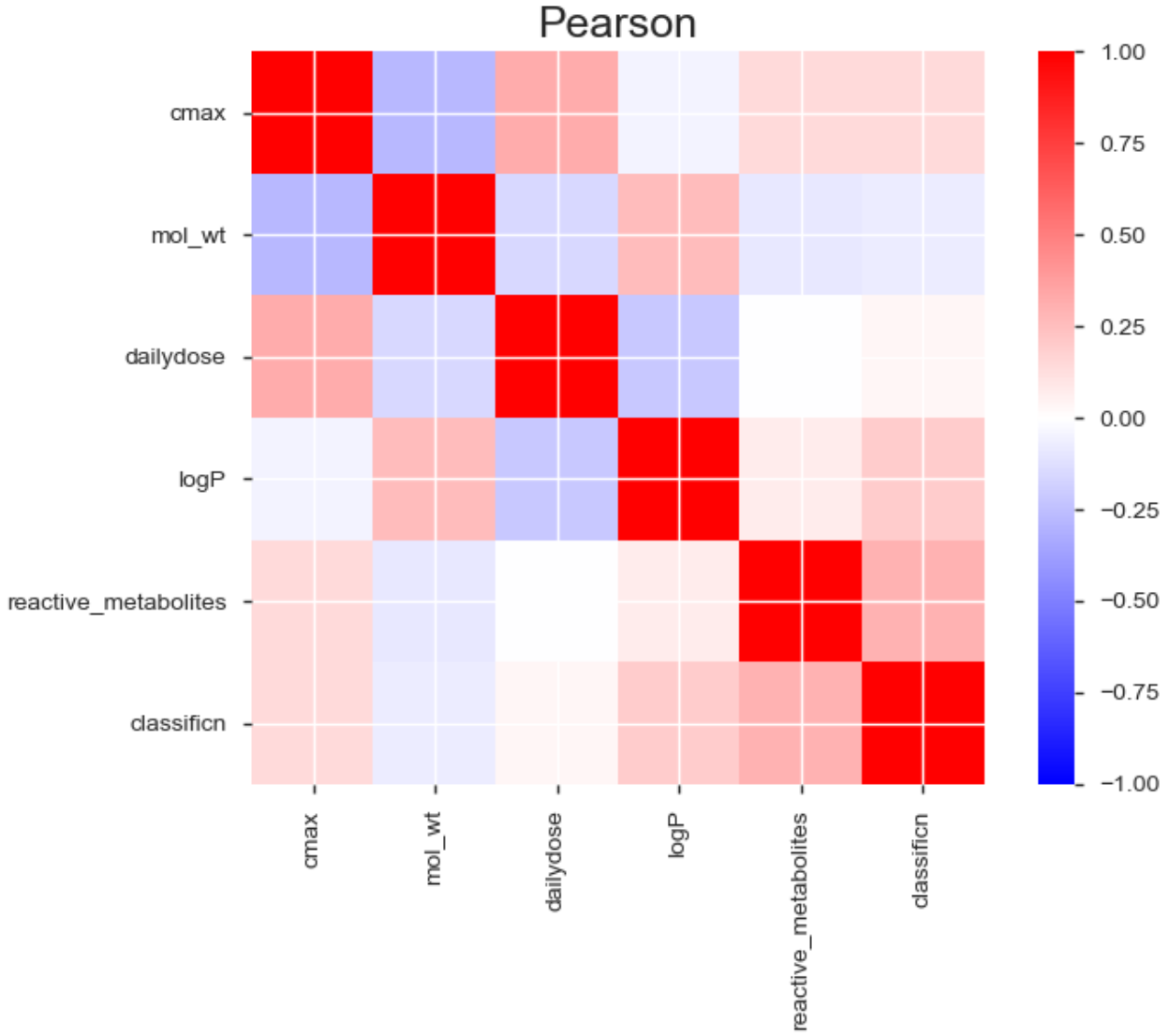


Data preparation for ML modeling

Generated dataset of 175 drugs consisting of:	• NAS values, corresponding Toxcast tests and their inferred (216) gene targets/pathways • DILI classification (60 “mostDILI”, 115 “other”) , drug/chemical name
Convert data to “wide” format:	• NAS values for all 216 Toxcast targets placed alongside each drug name • Missing NAS values imputed to -1000 (97% data consists of missing values)
Generate/visualize 3 “scaled”* datasets before modeling :	• Current predictors alone – “mc” data** • ToxCast predictors alone – “tc” data • Current predictors + ToxCast predictors – “mctc” data

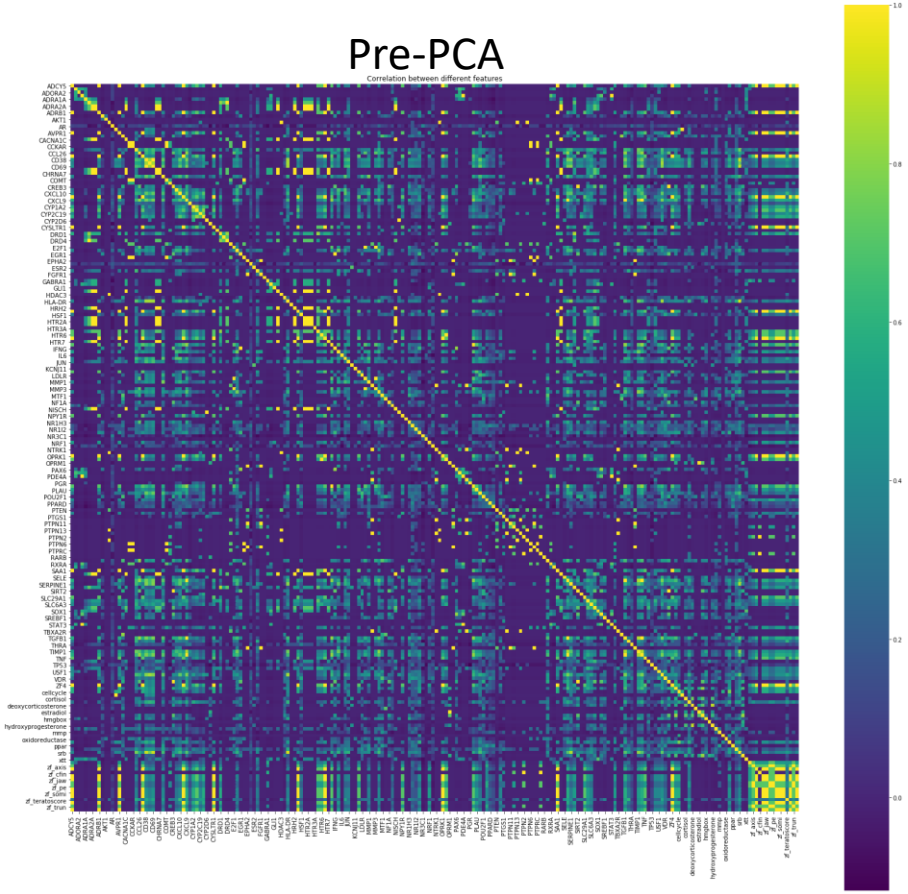
** Scaled using Standard Scaler in sklearn, **Current drug predictors consist of: hydrophobicity/logP, molecular weight, Cmax, daily dose, reactive metabolites (60 missing values for reactive metabolites imputed by literature search; all other values derived from LTKB and provided by Minjun Chen)*

Relationship between predictors in scaled “mc” data vs DILI classification (1=“mostDILI”, 0 = “other”)



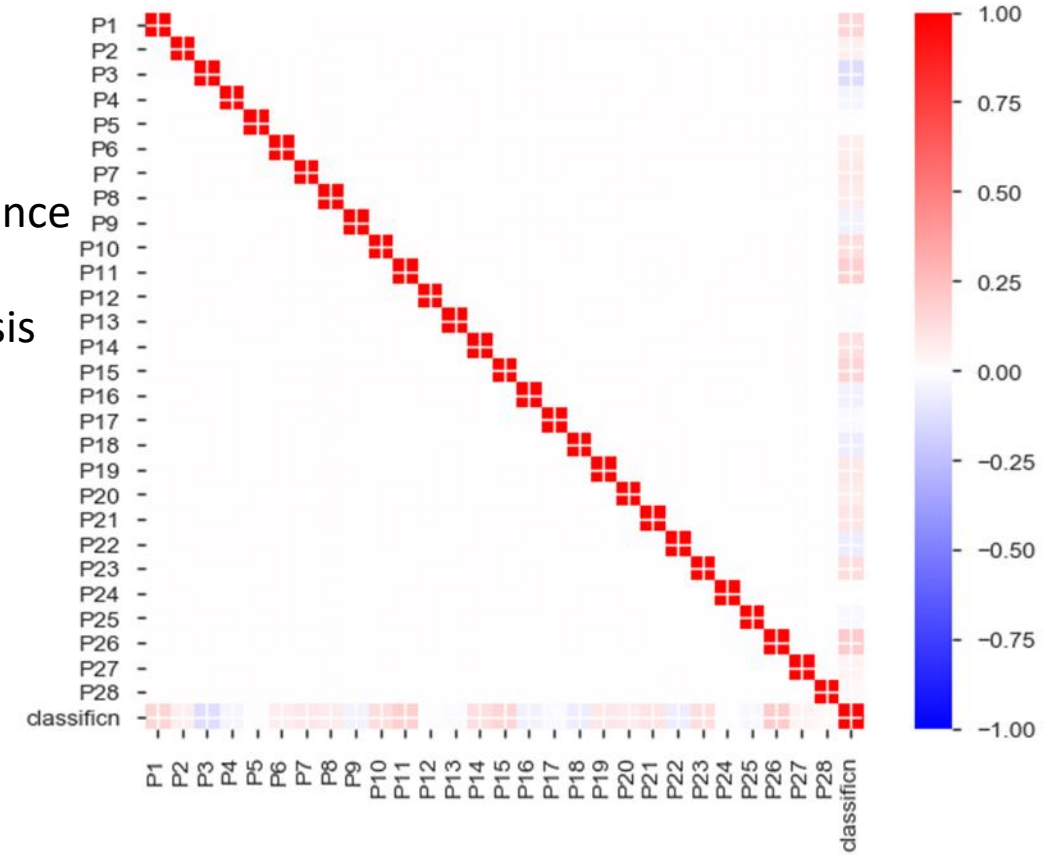
Relationship between ToxCast predictors vs DILI classification (pre-PCA vs post-PCA)

Pre-PCA



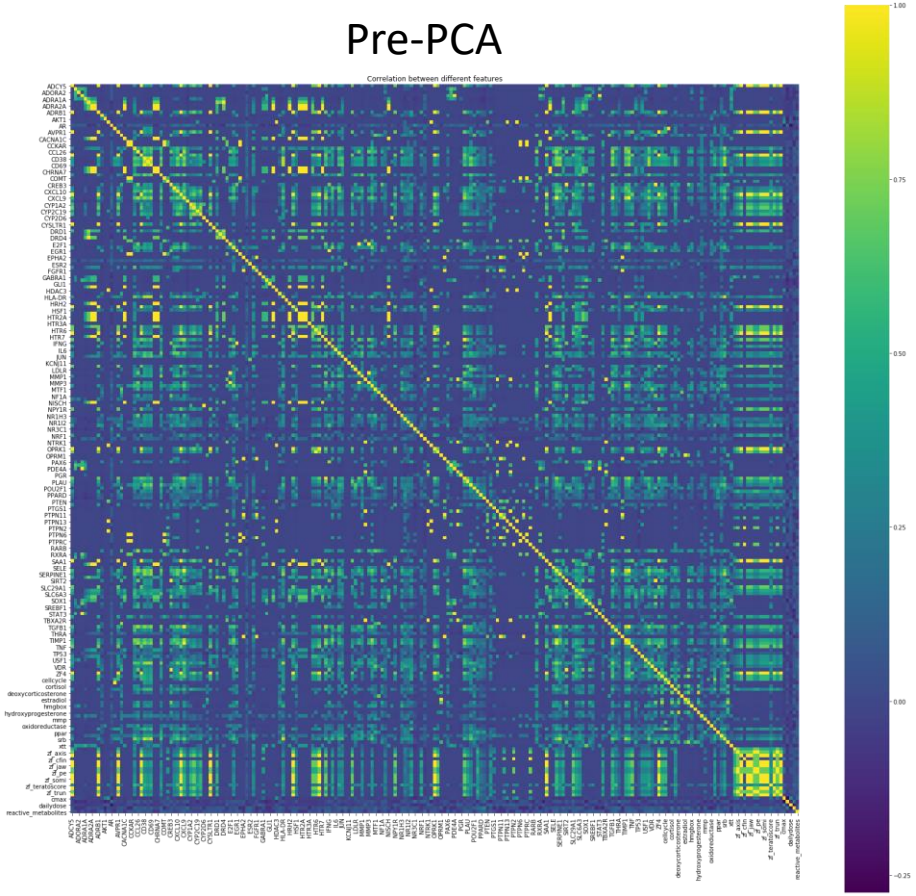
PCA @ 85% variance explained and correlation analysis

Post-PCA



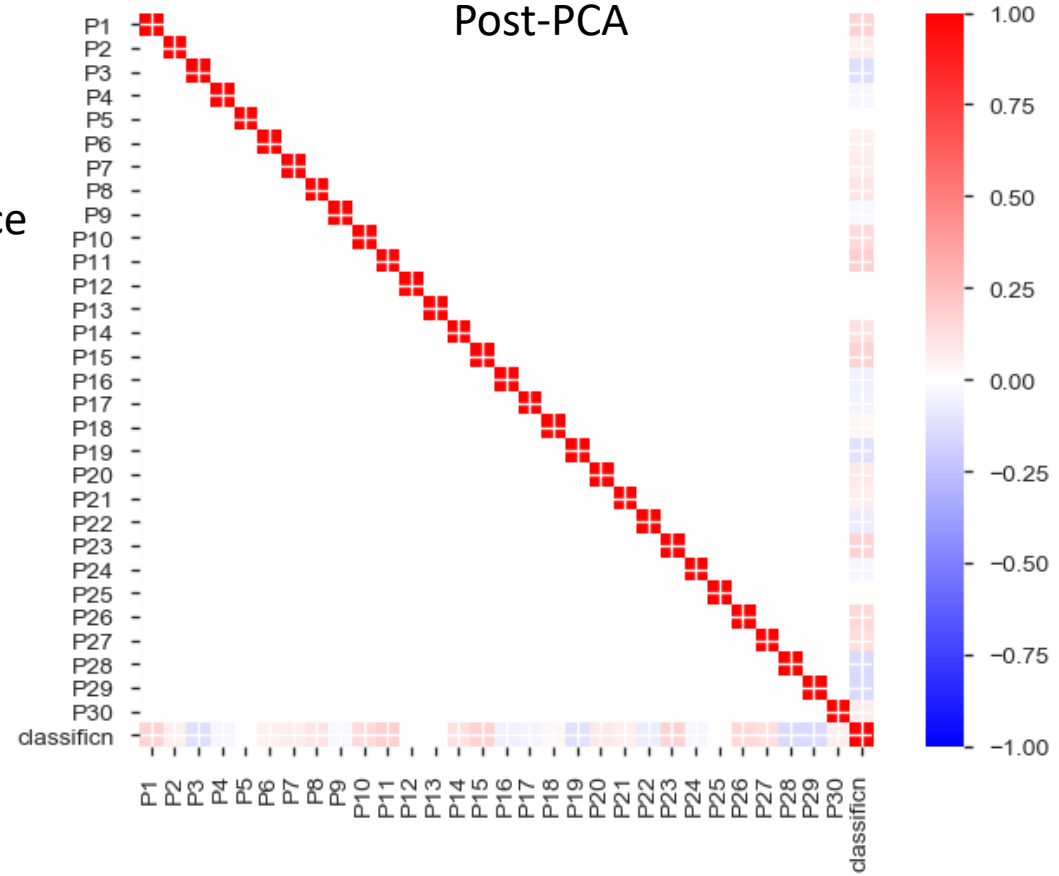
Relationship between all predictors and DILI classification (pre-PCA vs post-PCA)

Pre-PCA



PCA @ 85% variance explained and correlation analysis

Post-PCA



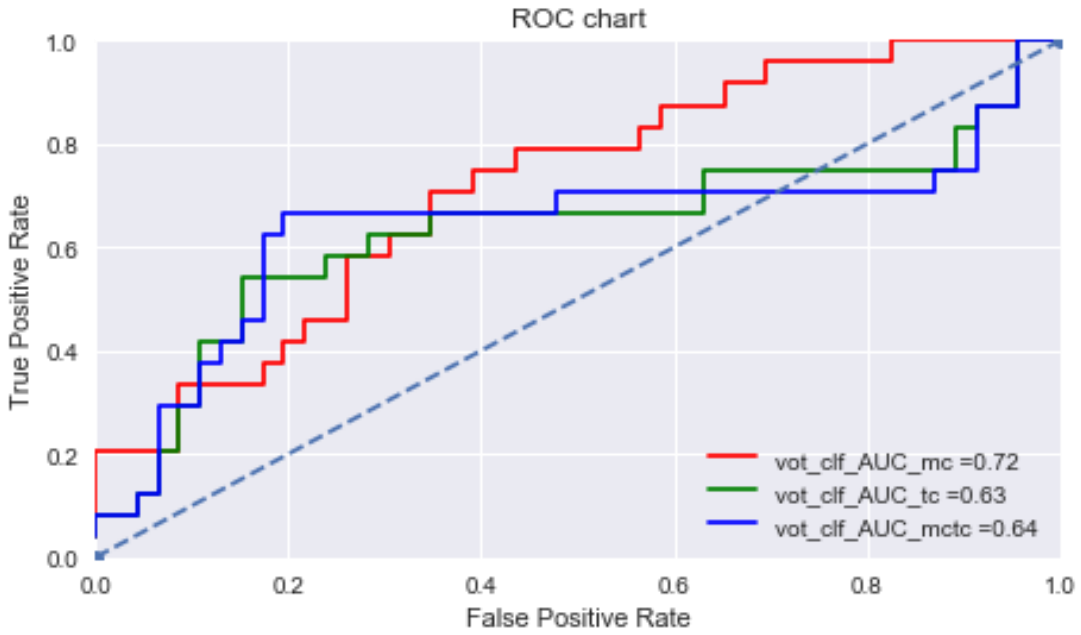
ML modeling strategy

Perform machine learning to predict “mostDILI” vs “other” DILI classification using:	• scaled mc data (benchmark model predictors) • post-PCA tc data (toxcast predictors) • post-PCA mctc data (benchmark model predictors + Toxcast predictors)
Evaluate performance of multiple machine learning algorithms*	• *Logistic regression*, trees/ensemble models (decision tree, random forest, gradient boost), KNN, naïve bayes, *support vector machine,* to be included • For all models maximize performance using hyperparameter tuning/grid-search CV and/or optimization of probability thresholds
Evaluate top (performing) algorithms over multiple test/train data combinations (10-fold cross validation)	• Combine best performing algorithms via voting classifier • Compare performance of top 2 performers with 10-fold CV for mctc data vs mc data alone across: *precision*, *sensitivity, specificity, accuracy*

* *All modeling performed in Python/sklearn package*

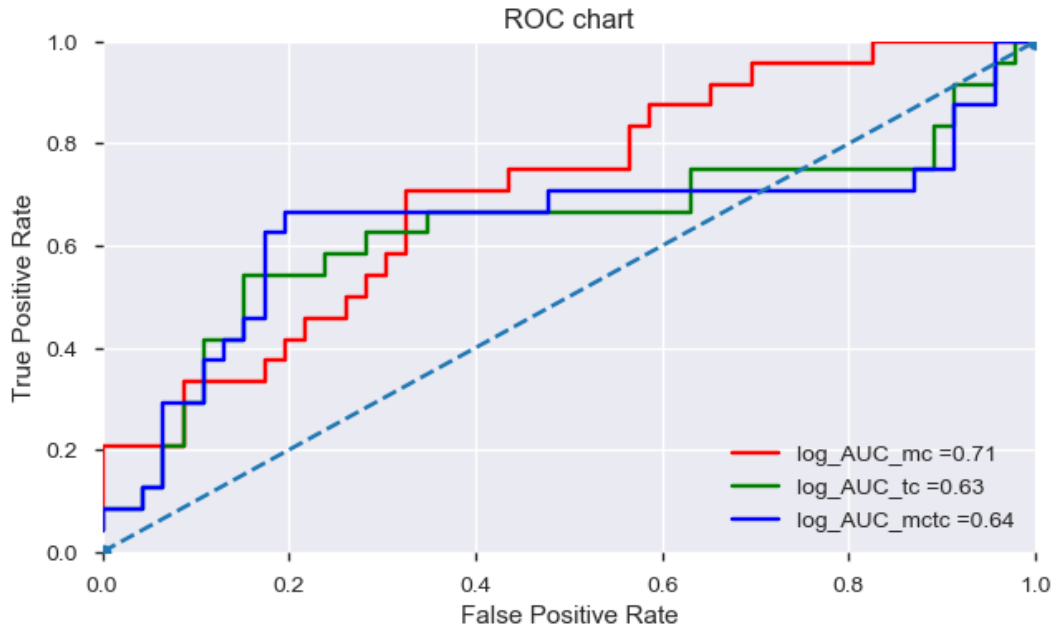
Best 2 models show additive effect of Toxcast predictors to “benchmark model/predictors”

“Fused”* logistic regression-SVM (best model)



“mctc” has modestly higher TPR relative to “mc” dataset but lower AUC across all probability thresholds

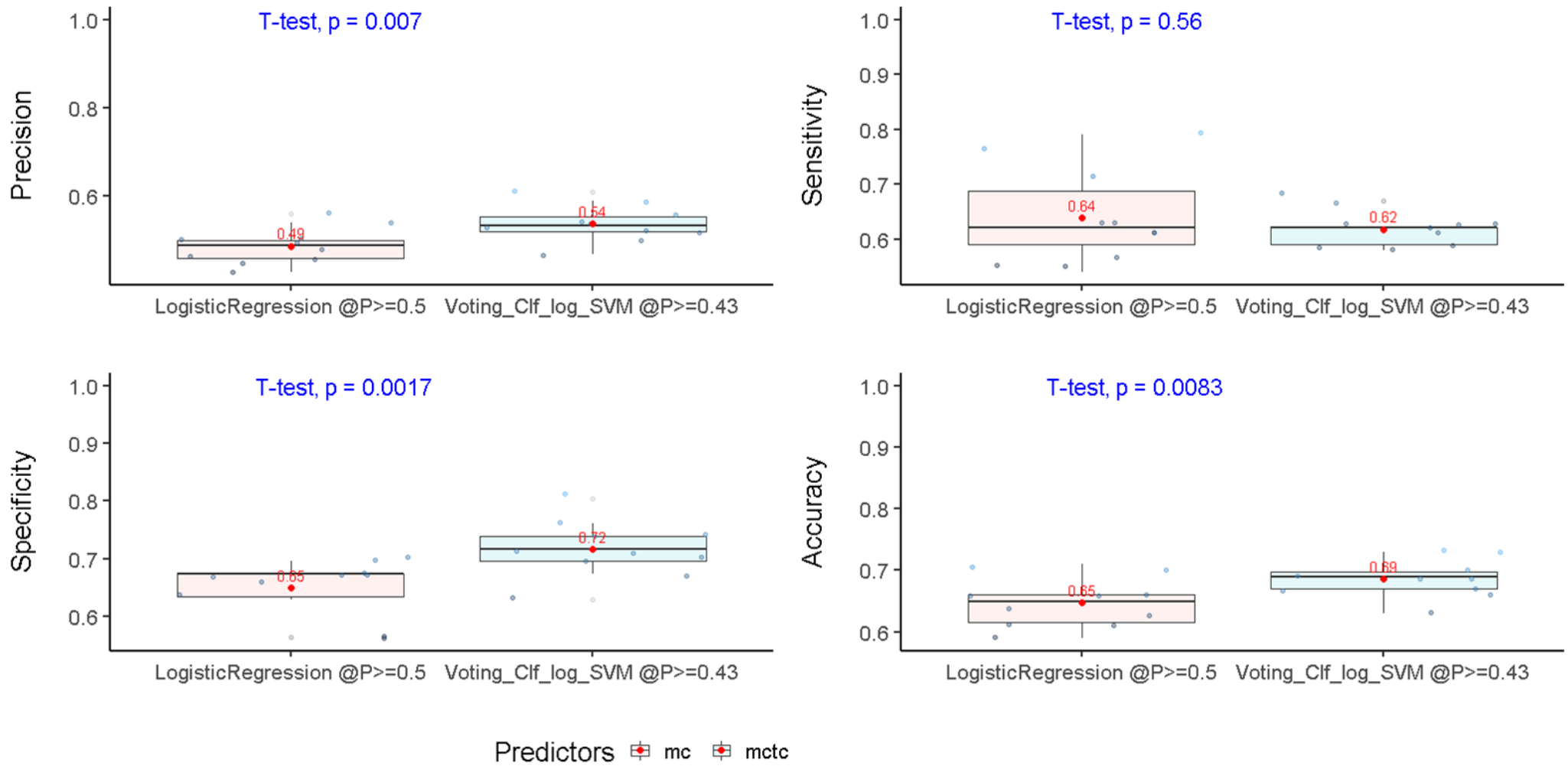
Logistic regression (next best model)



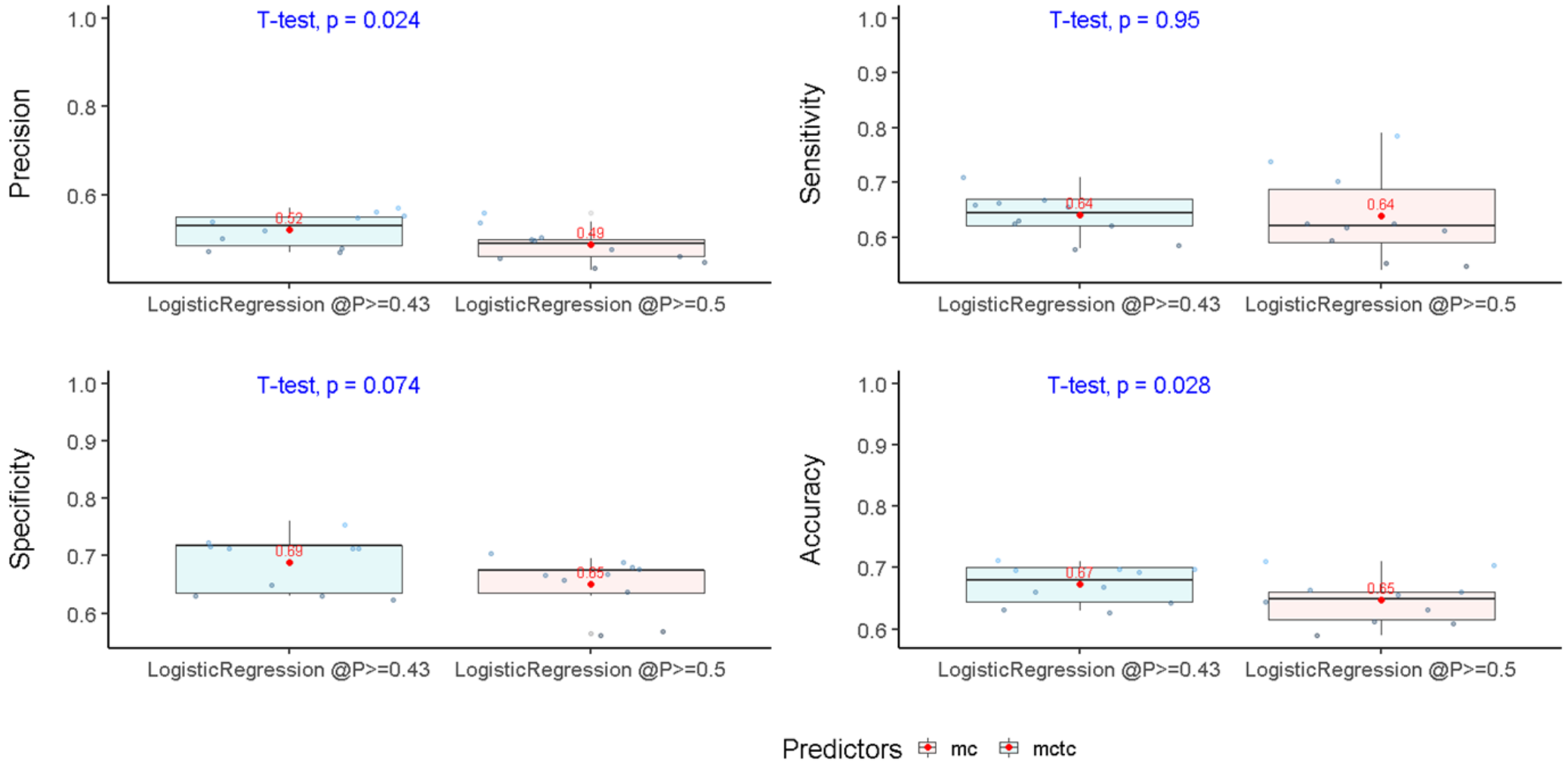
“mctc” has modestly higher TPR relative to “mc” dataset but lower AUC across all probability thresholds

*Voting classifier based on weighted averaging of probabilities from each algorithm used for “fusion”

Performance of fused logistic-SVM classifier (best model) with Toxcast predictors vs benchmark model



Logistic regression w/ optimized thresholds and Toxcast predictors vs benchmark model



Results summary

Toxcast predictors' addition enhances precision, specificity, and accuracy of DILI prediction models

	Model	Precision	Sensitivity	Specificity	Accuracy
Benchmark model	Logistic regression @P>=0.5	0.49	0.64	0.65	0.65
Alternative models with Toxcast predictors	Fused logistic-SVM @P>=0.43	0.54	0.62	0.72	0.69
	Logistic regression @P>=0.43	0.52	0.64	0.69	0.67

Value-add of Toxcast predictors to prediction of "mostDILI" drugs vs "other" DILI is quite modest

- *There is a need for alternative predictors for further enhancing DILI prediction models to be practically useful*

Limitations

1	**Reduction of Toxcast assay responses to gene targets/pathways**	Results in loss of “variance” from cell type, organism, time points.
2	**Imputation schema for targets with less than medium NAS**	All targets with < medium NAS AND *those without any data on perturbation* set to a “not activated” NAS level of -1000
3	**No mechanistic analysis of perturbed gene targets/pathways performed**	May be useful to “weight” predictors and improve performance
		Could not perform fold-change analysis of gene targets or pathways as alternative to AC50 values