Predicting *In Vivo* Toxicity from *ToxCast* Data Using Active Machine Learning

QUANTITATIVE MEDICINE
Transforming Drug Discovery™

The work discussed below was originally presented as a poster at the EPA ToxCast summit

To demonstrate the advantages and capabilities of the **Co**mputational **R**esearch **E**ngine™ (***CoRE***™) for predictive model building using Active Machine Learning, Quantitative Medicine and collaborators from a large pharmaceutical company undertook a two phase project for predictive modeling the ToxCast data and comparing to industry standard model building practices.

Enriched Discovery:

CoRE

Predictive Model Building: Learn predictive models using various machine learning methods (Lasso, SVM, RandomForest, etc.). For each endpoint, select the best model based on training set cross-validation accuracy.

CoRE™ **KnowledgeBase™:** The most relevant set of features calculated for each compound derived from a compilation of 200 million experimental results reduces redundant experiments.

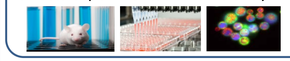
Model Generation

Active Machine Learning Cycle

Experiment Execution

Experiment Selection

Flexible/Robust: Highly accurate models across many types of assays directs experimentation efficiently.



Focuses Complex Research: Prioritizes compounds and experiments expected to most improve the predictive model towards the optimization goals of the project.

ToxCast *In Vitro Assays* *In Vivo* Endpoint Compound Data

Platform	Technology
ACEA	cell growth kinetics
AttaGene	reporter gene assay
BioSeek	ELISA assay
Cellumen	high content screening
CellzDirect	expression level change
GenTronix	growth arrest measure
NCGC	reporter gene assay
NovaScreen	enzyme inhibition assay
Solidus	microarray

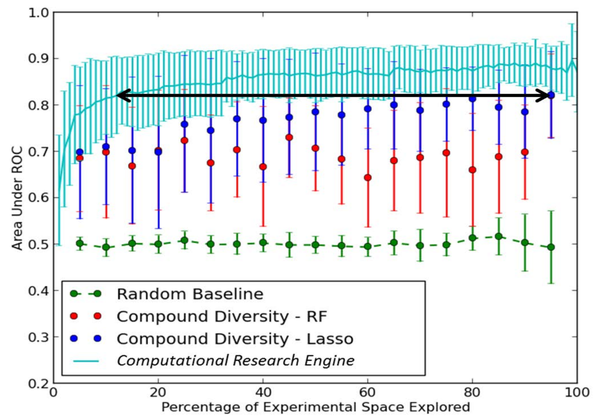
In Vivo Endpoint	Neg	Pos	% Pos
Systemic Necrosis	**146**	**167**	**53%**
Hepatocellular Necrosis	**230**	**115**	**33%**
Hepatocellular Necrosis + ALT⁺	**207**	**61**	**23%**
Hepatocellular Necrosis + ALT⁺ + Bilirubin⁺	**146**	**17**	**10%**
Liver Injury	**67**	**224**	**77%**
Kidney Tubule Toxicant	**310**	**52**	**14%**
Bile Duct Hyperplasia	**225**	**56**	**22%**

***CoRE's* capabilities are demonstrated** across 887 *in vitro* assays simultaneously to direct experimentation, and across 7 *in vivo* toxicity endpoints simultaneously to leverage data from the *KnowledgeBase*.

Phase I – *In Vitro* Assay Cyclical Predictive Model Training

ActiveLearning™ vs Diversity-Paired Machine Learning
309 cmpds, 887 *in vitro* assays – Incremental Data Build

(simulation protocol details on back)



Metric	CoRE™	Lasso	RandomForest	Random
AU-ROC(5% Data)	0.77	0.7	0.68	0.5
AU-ROC(10% Data)	0.82	0.71	0.7	0.48
Max AU-ROC	0.89	0.83	0.73	0.52
% Data(Max AU-ROC)	99%	95%	45%	85%
Prediction Accuracy Curve Description (% Data)	rapid gain 1-5%, mod. gain 5-40%, 55-95%	plateau 5-20%, sinusoidal gain 25-95%	unstable over simulation	stable monotonic at random

- The ***CoRE*™ guided data selections** increased predictive model accuracy significantly in the first 5% data explored (4 cycles) and continued improvement, reaching > 0.8 by 8% (7th cycle) and a maximum of 0.89.
- The **Diversity-Lasso paired** simulation reached 0.75 at 25% of data explored and a maximum of 0.83.
- The **Diversity-RandomForest paired** simulation never reached > 0.73 and the models did not stabilize.
- The **Random** selection models performed as expected.

CoRE's directed *ActiveLearning*™ approach results in highly predictive models which **explore only 1/5th** the amount of data space compared to the next best standard method.

In a discovery cycle campaign, *CoRE*™ can effect significant savings in time, resources and investment.

Phase II – *In Vivo* Toxicity Models from Leveraged Data

KnowledgeBase™ signal vs *ToxCast in vitro* assay signal
648 cmpds, 887 *in vitro* assays, 7 *in vivo* tox endpoints

(simulation protocol details on back)

AU-ROC curve values for various feature sets

In Vivo Toxicity Endpoint	SFP	CoRE	ToxCast	All
Systemic Necrosis	0.66	0.61	0.51	0.66
Hepatocellular Necrosis	0.69	0.75	0.6	0.76
Hepatocellular Necrosis + ALT+	0.7	0.76	0.74	0.79
Hepatocellular Necrosis + ALT+ + Bilirubin+	0.58	0.78	0.5	0.66
Liver Injury	0.67	0.71	0.59	0.58
Kidney Tubule Toxicant	0.5	0.6	0.49	0.59
Bile Duct Hyperplasia	0.68	0.72	0.68	0.82
Mean	0.64	0.7	0.59	0.7

- *CoRE*'s use of its *KnowledgeBase* yields higher accuracy models than models from *ToxCast in vitro* results in all endpoints tested.
- Addition of *ToxCast* data (all) can lead to worse models.
- The best models contain significant amounts of data from the *KnowledgeBase* (CoRE_Feature).

In Vivo Toxicity	Feature Name	Feature Relevance
HEPAT_NECROSIS	CoRE_Feature_Screen_Anti-Amyloidogenic_Compounds	0.96
	CoRE_Feature_Inhib_T4bacteriophage_DNA polymerase	0.82
	TxCA_Feature_APR_CellCycleArrest_1h_up	0.81
	TxCA_Feature_APR_CellCycleArrest_24h_dn	0.70
	TxCA_Feature_ACEA_T47D_80hr_Negative	0.52
HEPAT_NECROSIS_ALT	CoRE_Feature_Screen_Anti-Amyloidogenic_Compounds	0.94
	CoRE_Feature_Inhib_T4bacteriophage_DNA polymerase	0.77
	TxCA_Feature_APR_CellCycleArrest_1h_up	0.66
	TxCA_Feature_APR_CellCycleArrest_24h_dn	0.55
	TxCA_Feature_NVS_LGIC_rGluNMDA_Agonist	0.41
HEPAT_NECROSIS_ALT_BILI	CoRE_Feature_Screen_Inhib_ClassIIHMG-CoAReductases	1.00*
	CoRE_Feature_Inhib_SUMOylation	1.00*
	CoRE_Feature_Inhib_Ratbrain_Monoacylglycerol_Lipase	1.00*
SYSTEMIC_NECROSIS	CoRE_Feature_Inhib_T4bacteriophage_DNA polymerase	0.00**
	CoRE_Feature_Screen_Anti-Amyloidogenic_Compounds	0.00**
	CoRE_Feature_Screen_ABC_Transporter_Inhib	0.00**
	CoRE_Feature_Screen_Mcl-1/Noxa_Interaction_Inhib	0.00**
	CoRE_Feature_Screen_Compounds_Affecting_Thrombin_Function	0.00**

The Top 5 Most Informative Features for Highly Predictive Models

The **advantage of Quantitative Medicine's *CoRE*** method was amply demonstrated in a two phase project for predictive modeling of the *ToxCast in vitro* and *in vivo* data.

- Phase I clearly demonstrated ***ActiveLearning*** **more rapidly learns predictive models of higher accuracy** as compared to standard diversity based methods.
- Phase II highlighted the ***KnowledgeBase's*** **power to leverage data and yield models with greater signal** than models built from the *ToxCast* assay data alone.

Contact: Josh D. Kangas PhD, Chief Scientific Officer josh.kangas@qtmed.com www.QtMed.com +1.412.616.0641

Modeling Simulation Protocols

This work was originally presented as a poster at the EPA ToxCast summit in September 2014:
http://qtmed.com/wp-content/uploads/2014/09/ToxCast_Data_Summit_2014_Final.pdf

Phase I Simulation Protocol – *In Vitro* Assay Cyclical Predictive Model Training

ActiveLearning™ vs Diversity-Paired Machine Learning
309 compounds, 887 *in vitro* assays – Incremental Cyclical Data Build

Starting from an initial random pool of *ToxCast in vitro* assay data for model training, and then incrementing the training set by two strategies, predictive models were built using *CoRE*, Lasso and RandomForest machine learning approaches to assess the relative prediction accuracy of the methods as a function of the total data explored.

- *CoRE* is an Active Machine Learning approach which utilizes additional data leveraged from its *KnowledgeBase*.
- Lasso is a regularized regression machine learning technique that avoids over-fitting of training data and is useful for feature selection.
- Random Forest is a classification machine learning algorithm that involves generation of an ensemble of randomly generated decision trees.

(Random Baseline: As a control, random experiments were selected and given random predictions to verify prediction accuracy calculations. Accuracy values of 0.5 are expected.)

After predictive models were generated, predictions on the remainder of the data set were made and the accuracy of each method was calculated. This iterative process was repeated 20 times from different starting points for each method to determine the mean and standard deviation of prediction accuracies.

In the next cycle iteration, the amount of data used for training was incremented (by 1% for *CoRE* and 5% for Lasso and Random Forest). *CoRE's* method for guided selection was compared to a standard diversity based selection. Diversity based selection is commonly used to select compounds for a model building task. The most diverse n% of compounds were found by clustering compounds using their fingerprints. Using K-Means, the compounds were clustered into as many clusters as compounds we desired to select and then a random compound was chosen from each cluster.

This process was repeated for incremental amounts of train/test set for 100% of the data.

In the figure we plot the mean and standard deviations for all train/test set increments for each method to visually assess the relative prediction accuracy of the methods. In the table we capture key features of the figure that distinguish the compared methods.

Phase II Simulation Protocol – *In Vivo* Toxicity Models from Leveraged Data

KnowledgeBase™ signal vs ToxCast *in vitro* assay signal
648 compounds, 887 *in vitro* assays, 7 *in vivo* toxicity endpoints – Static Models

Starting from a random selection of 75% train/25% test, predictive models were built using various feature sets and their combinations for *in vivo* toxicity endpoints to assess the relative strength of information captured by each method and the resulting prediction accuracy (AU-ROC) of derived models.

	Model Feature Sets
SFP	2D and 3D Structural Finger prints
CoRE	CoRE KnowledgeBase + SFP
ToxCast	ToxCast *in vitro assay* data + SFP
ALL	All features (SFP + CoRE + ToxCast)

After predictive models were generated, predictions on the test set were made and the accuracy (AU-ROC) of each feature set derived model was calculated.

Additionally, we ranked and tabulated the calculated feature relevance of the most highly predictive models for each endpoint using the All feature set to ascertain their source (*CoRE* or *ToxCast*) and their relative importance. The top 5 most relevant features are listed for the best predictive model. A feature relevance of 0.00 or 1.00 denotes a modeling method that does not allow a relative feature relevance to be calculated (ex. Lasso regression).

For reference, the following pathological observations were included as resulting in a positive classification for the in vivo toxicity endpoint.

In Vivo Endpoint	Included Pathological Observations
Liver Injury	**Necrosis, apoptosis, degeneration, vacuolization, regeneration**
Kidney Tubule Toxicant	**Necrosis, apoptosis, degeneration, vacuolization, regeneration, dilation**
Systemic Necrosis	**Internal organs excluding GI tract**

***CoRE* is a better lens for discovery, enabling research on your own terms**

Concurrent Optimization Enables a Clearer View (better predictions) and a Deeper Understanding (enriched knowledge)

Contact: Josh D. Kangas PhD, Chief Scientific Officer josh.kangas@qtmed.com www.QtMed.com +1.412.616.0641