



# QUANTITATIVE MEDICINE

Transforming Drug Discovery™

Technology Evaluation Consortium  
Cambridge Healthtech Associates

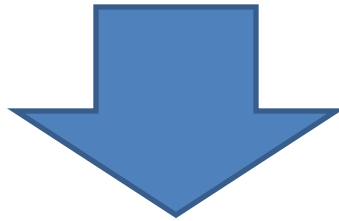
October 9, 2013

# This Presentation Contains Proprietary and Confidential Information

It may not be shared with any third party or be used for any purpose other than to evaluate a potential business relationship and/or perform a TEC study with Quantitative Medicine.

Copyright © 2012-2013. All Rights Reserved.  
*Quantitative Medicine, LLC*

**Analytics Focused Research System (AFRS)**



**Computational Research Engine (CoRE)**

# Agenda

- Hepatotoxicity Study Design
- Simulation Results
- Next Study

# Hepatotoxicity Study

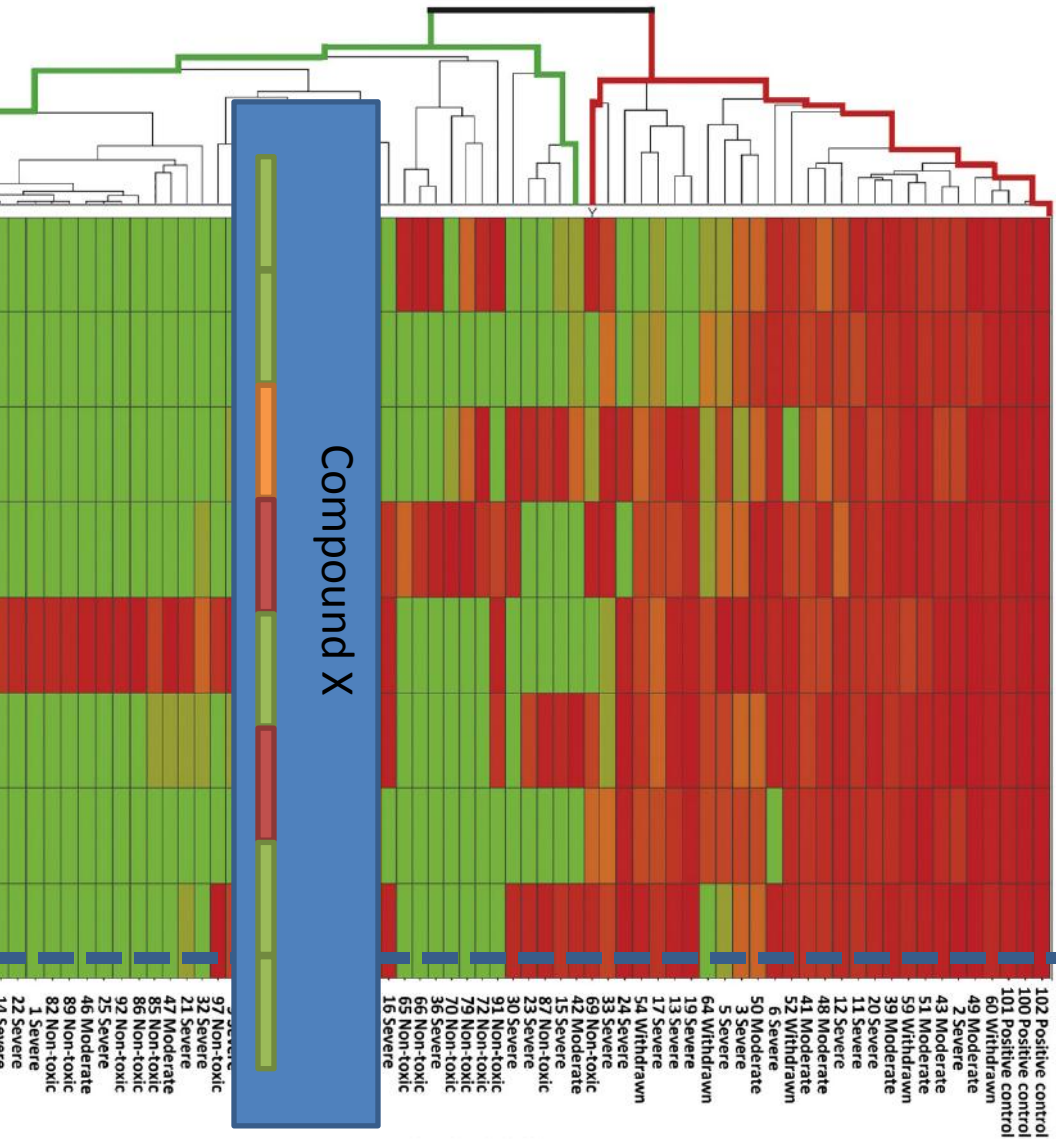
**Problem:** We want to accurately predict toxicity for new compounds based on the results of HCS experiments.

**Approach:** Test each new compound in HCS assay. Use the HCS results as inputs for a predictive model learned from prior experimentation.

8 HCS Measurements

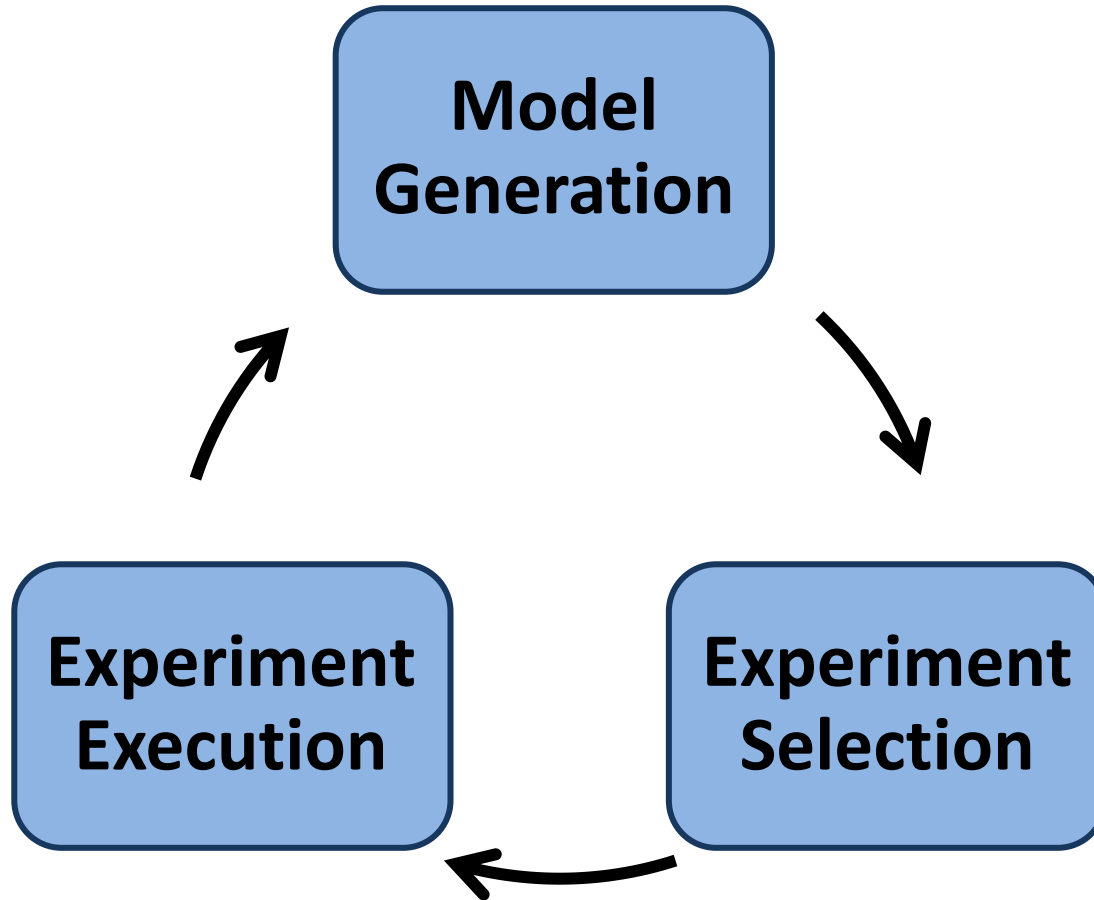
Overall Goal: Learn a model to *accurately* predict toxicity for a compound

Task: Given HCS observations, predict toxicity for Compound X.



105 Compounds

# Active Learning



# CoRE Simulation Protocol

- Dataset randomly divided into training set (80%) and testing sets (20%)
- All results (Toxicity and HCS measures) were hidden from CoRE
- For each round a compound was selected from the training set and the toxicity and HCS readouts were revealed to the learner
- A predictive model was learned from only the *revealed* training data
- Predictions were made for all compounds in the testing set



# Receiver Operating Characteristic

- What are we predicting?
  - Hepatotoxicity for all “unobserved” experiments in testing set
- How do we measure accuracy of those predictions?
  - Calculate a ROC curve by sweeping across prediction thresholds and comparing to hidden ground truth toxicity

# Comparison Method

**Purpose:** Develop a Model to Actively Learn and Accurately Predict Toxicity From HCS Data

**Standard Approach (single pass):**

**Selection method:** Choose a set of compounds randomly for “execution” from training set

**Prediction method:** Learn a RandomForest model based on HCS results to predict toxicity.

# Comparison Method

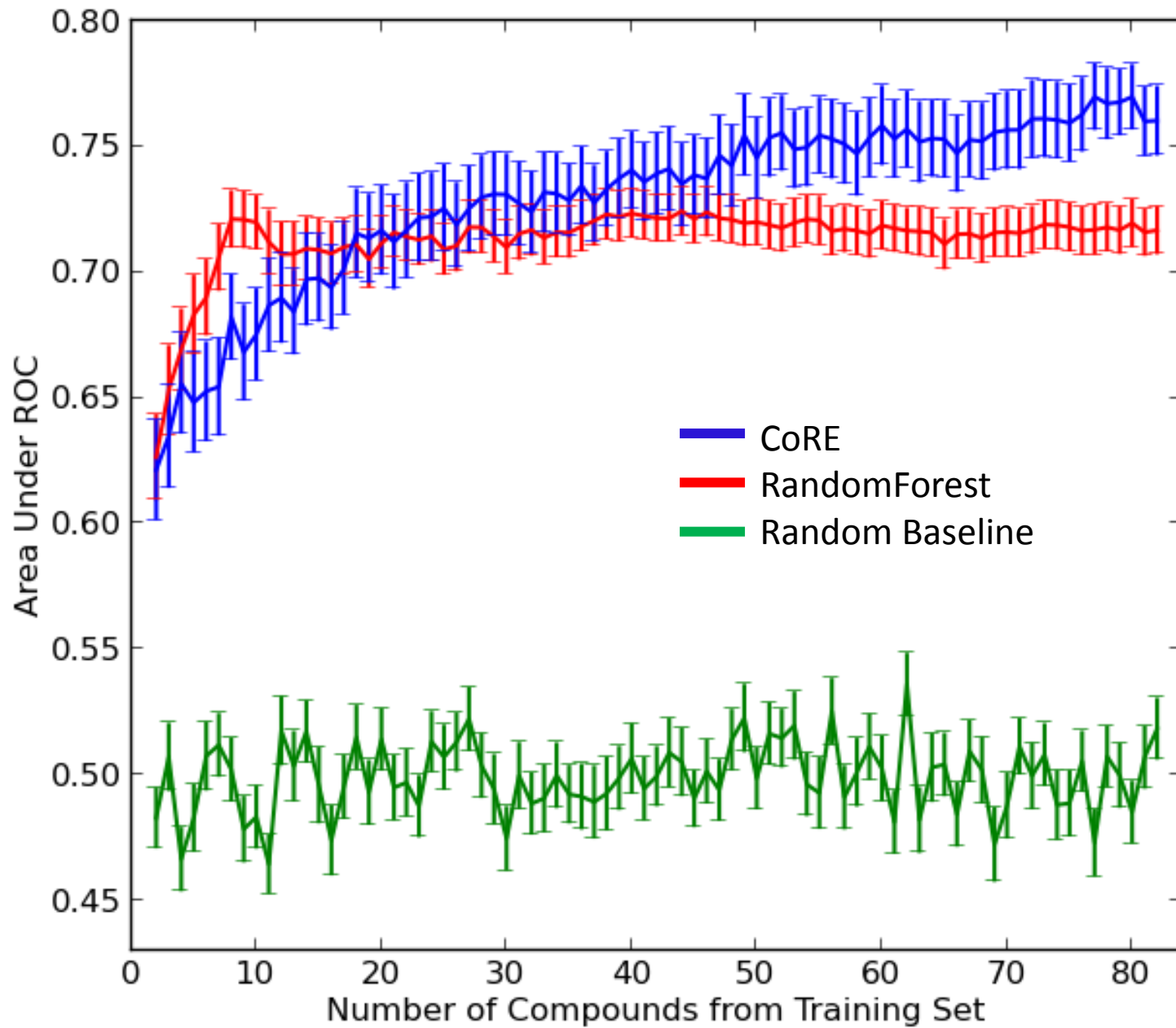
**Purpose:** Develop a Model to Actively Learn and Accurately Predict Toxicity From HCS Data

**CoRE Approach (iterative):**

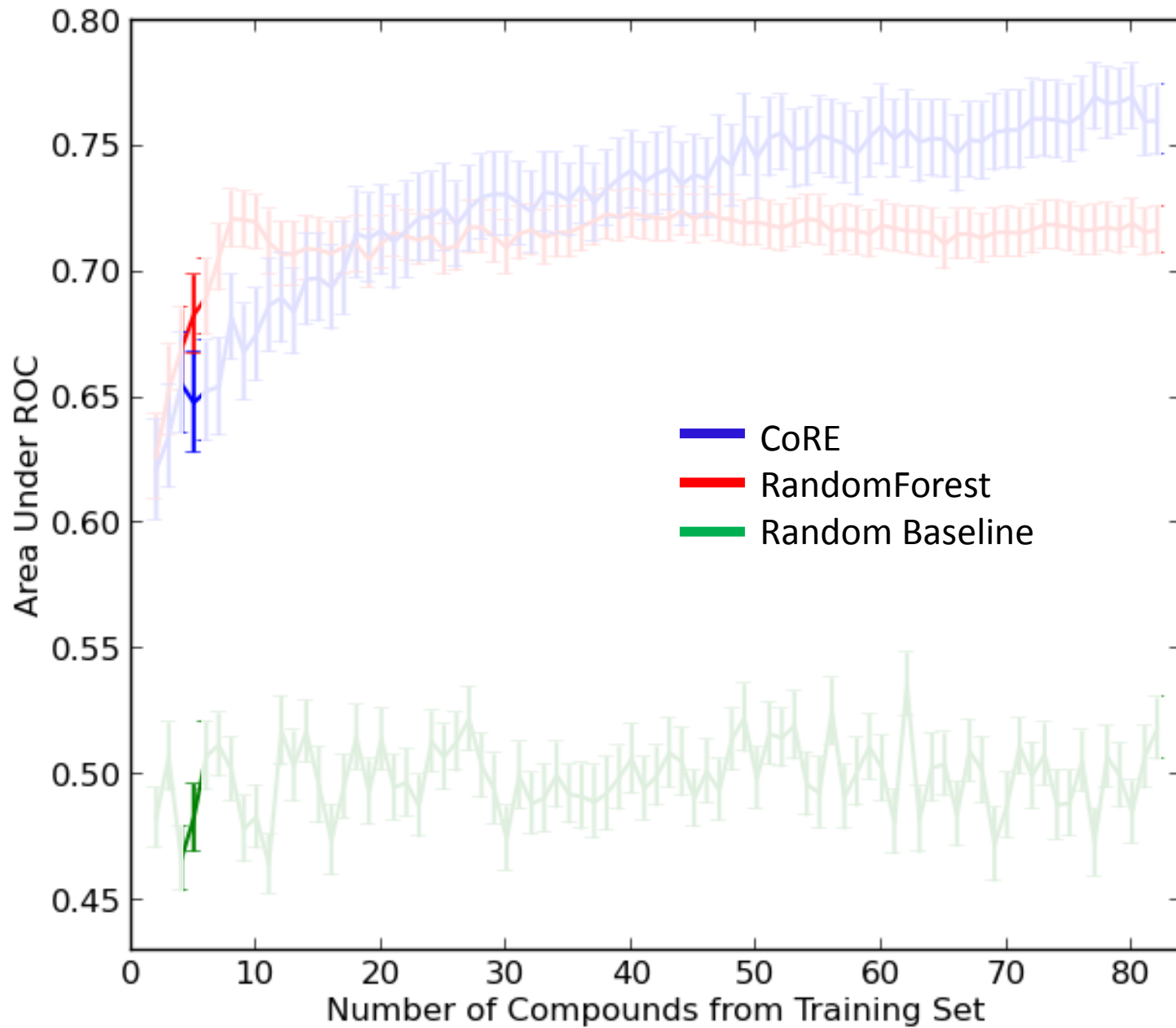
**Selection method:** Choose a single compound for “execution” from training set.

**Prediction method:** Learn a model using our prediction methods based on all revealed HCS results and our extensive library to predict toxicity.

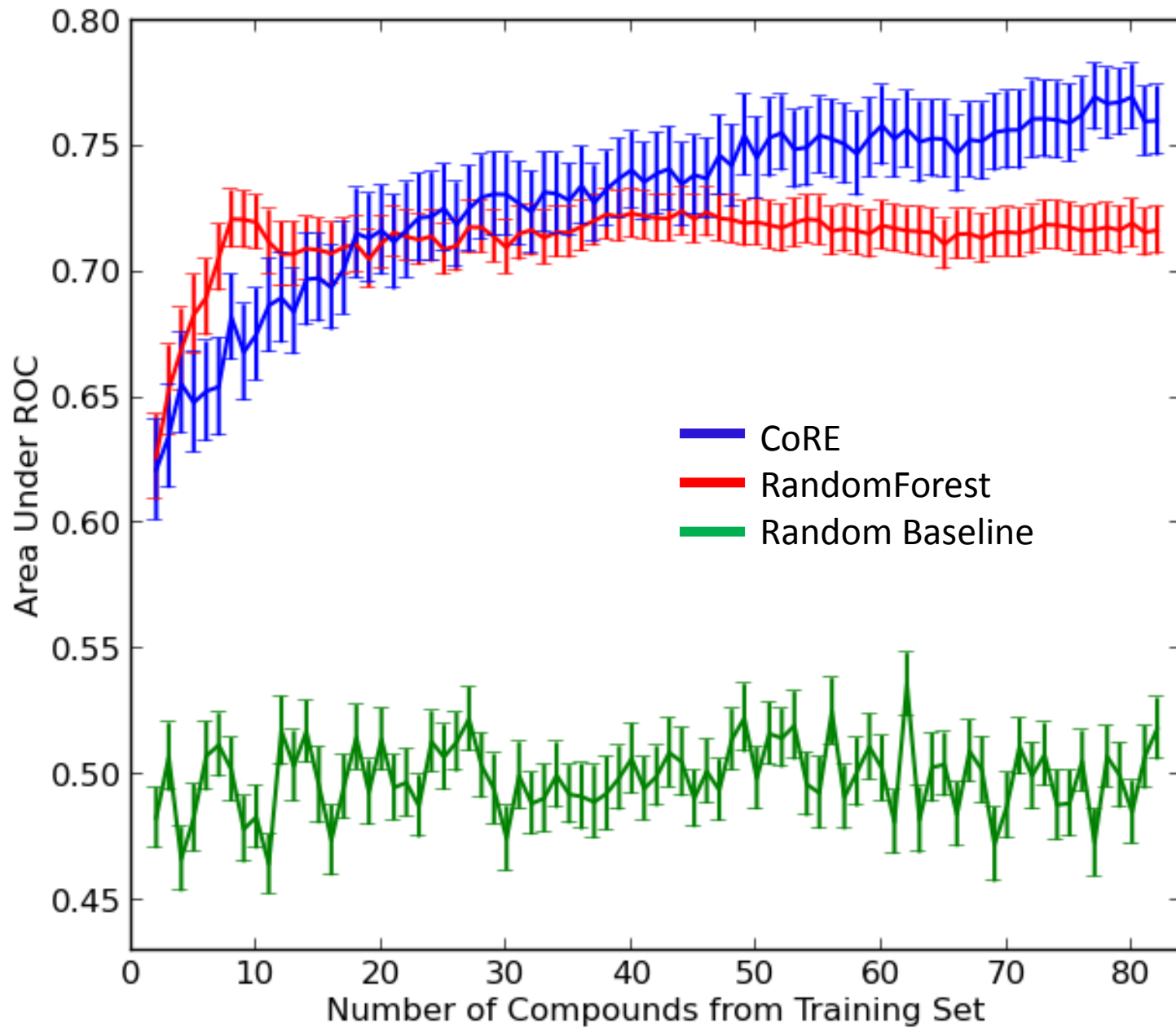
# Hepatotoxicity Simulation Results



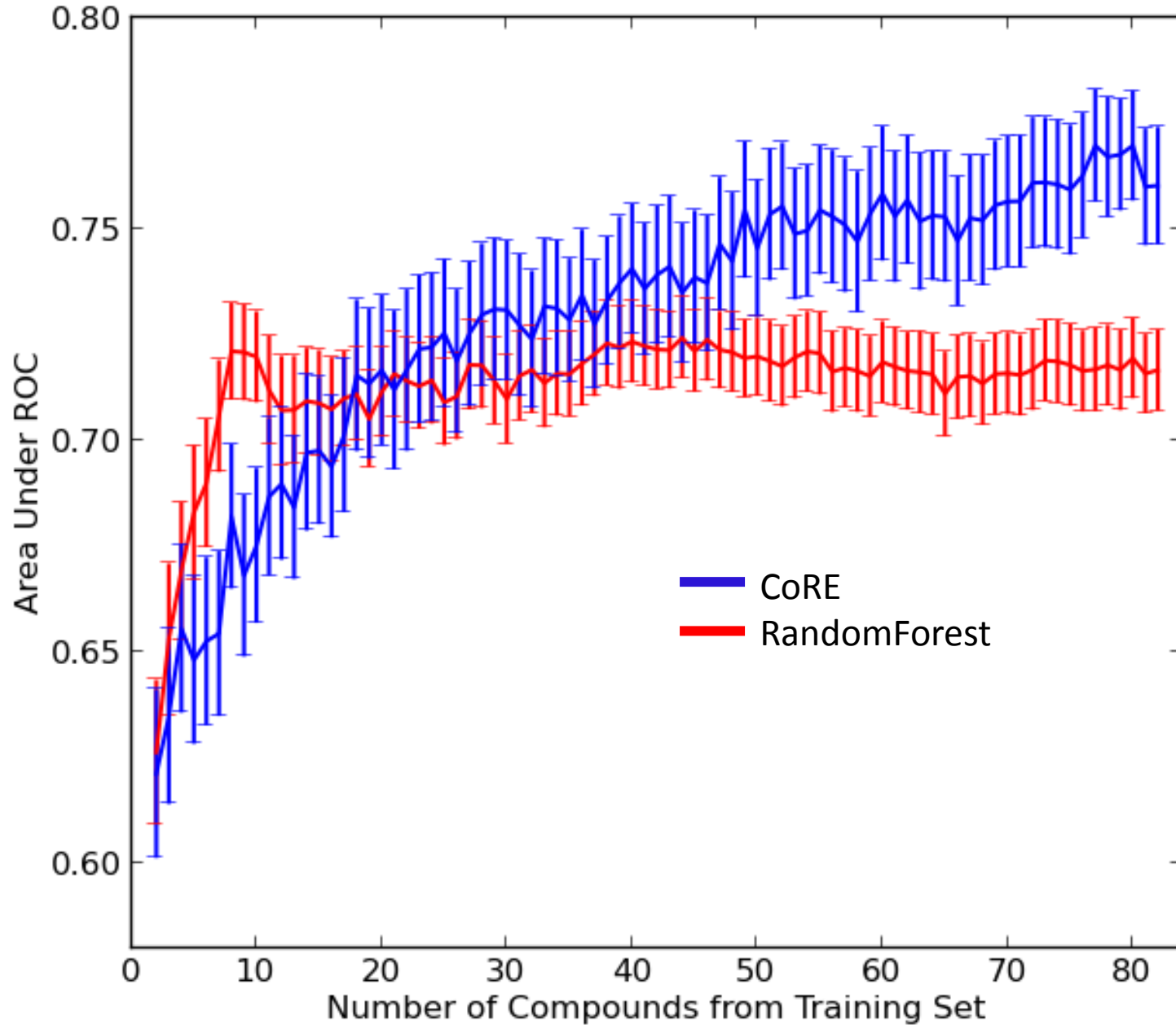
# Hepatotoxicity Simulation Results



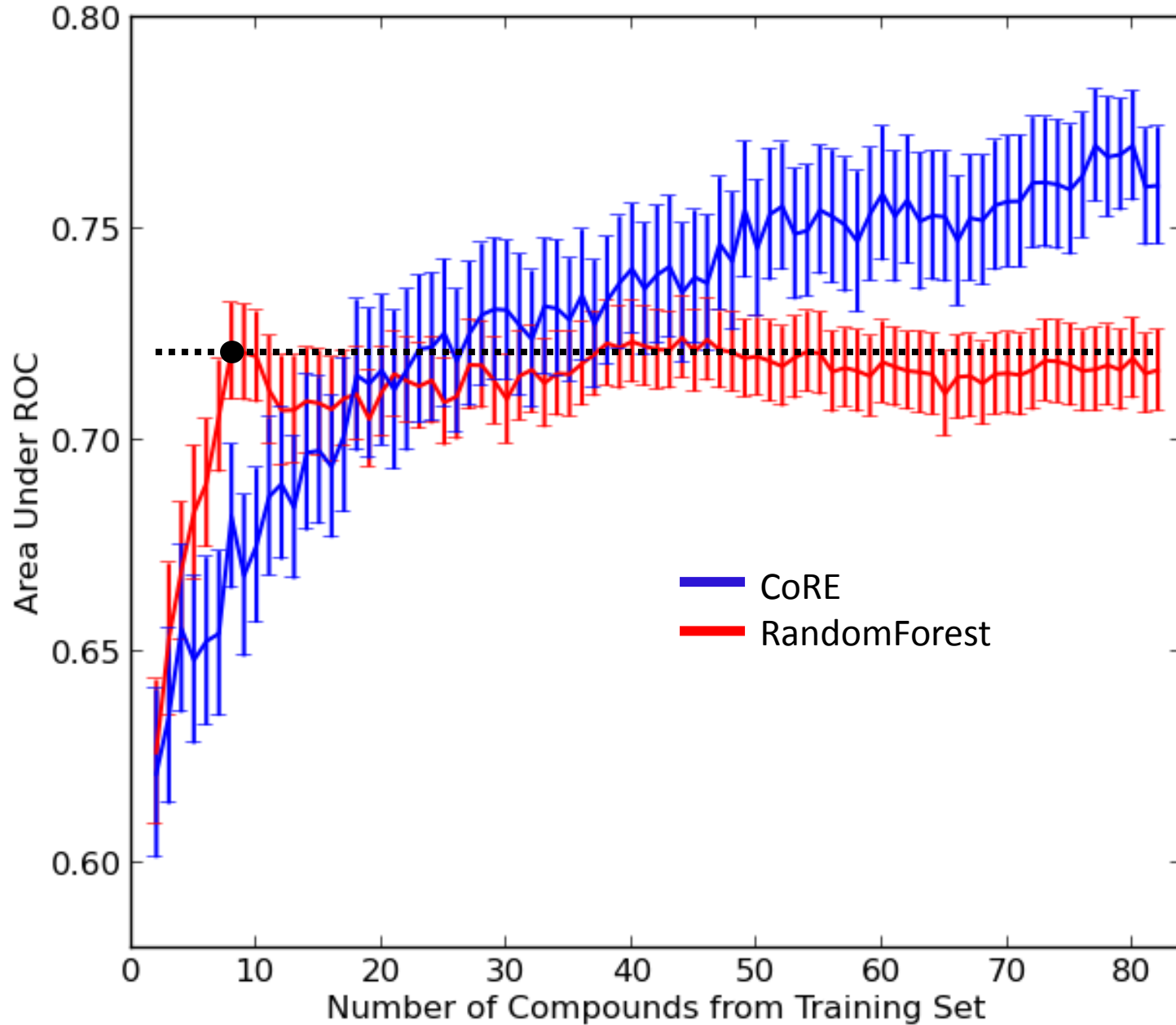
# Hepatotoxicity Simulation Results



# Hepatotoxicity Simulation Results

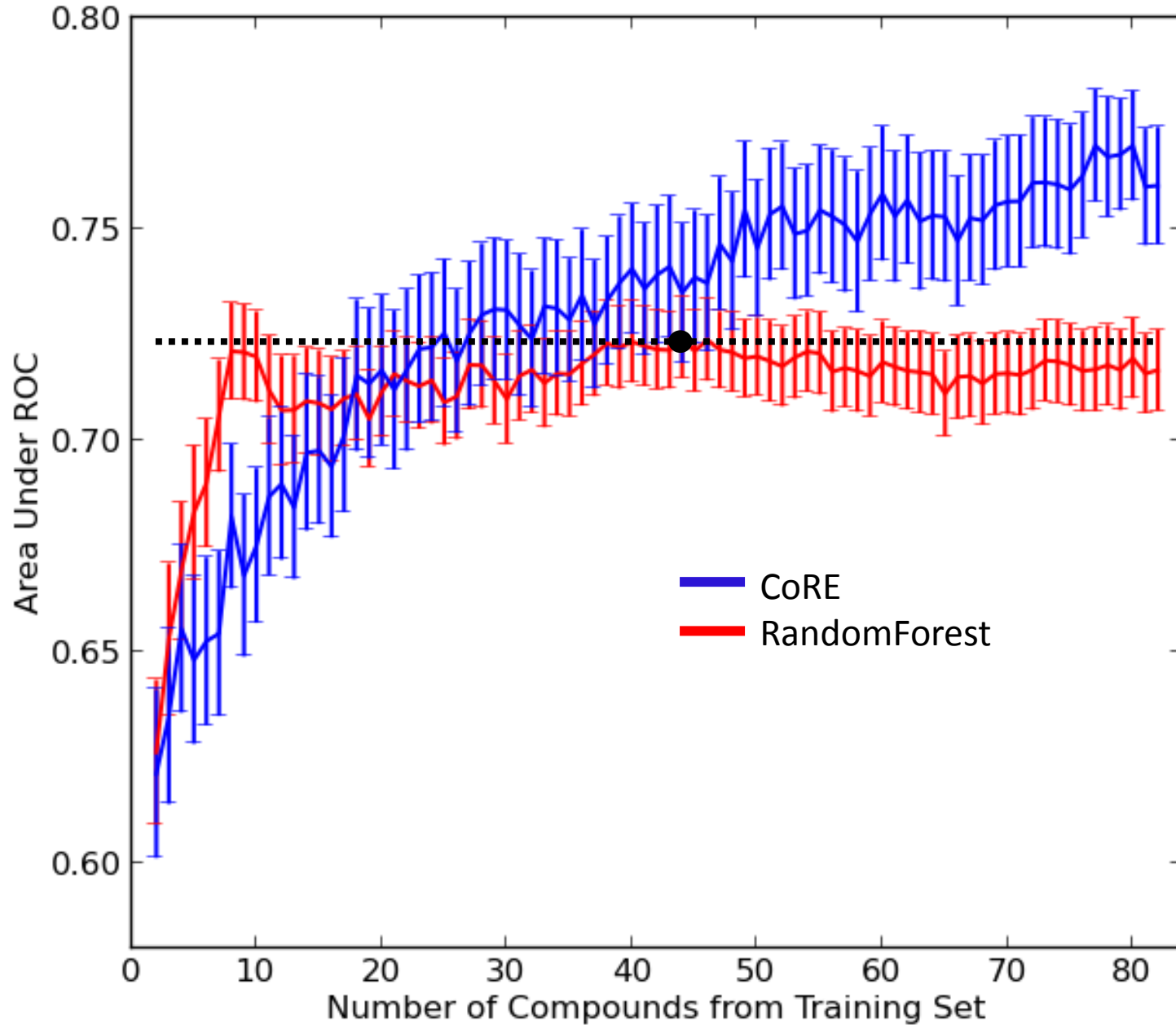


# Hepatotoxicity Simulation Results

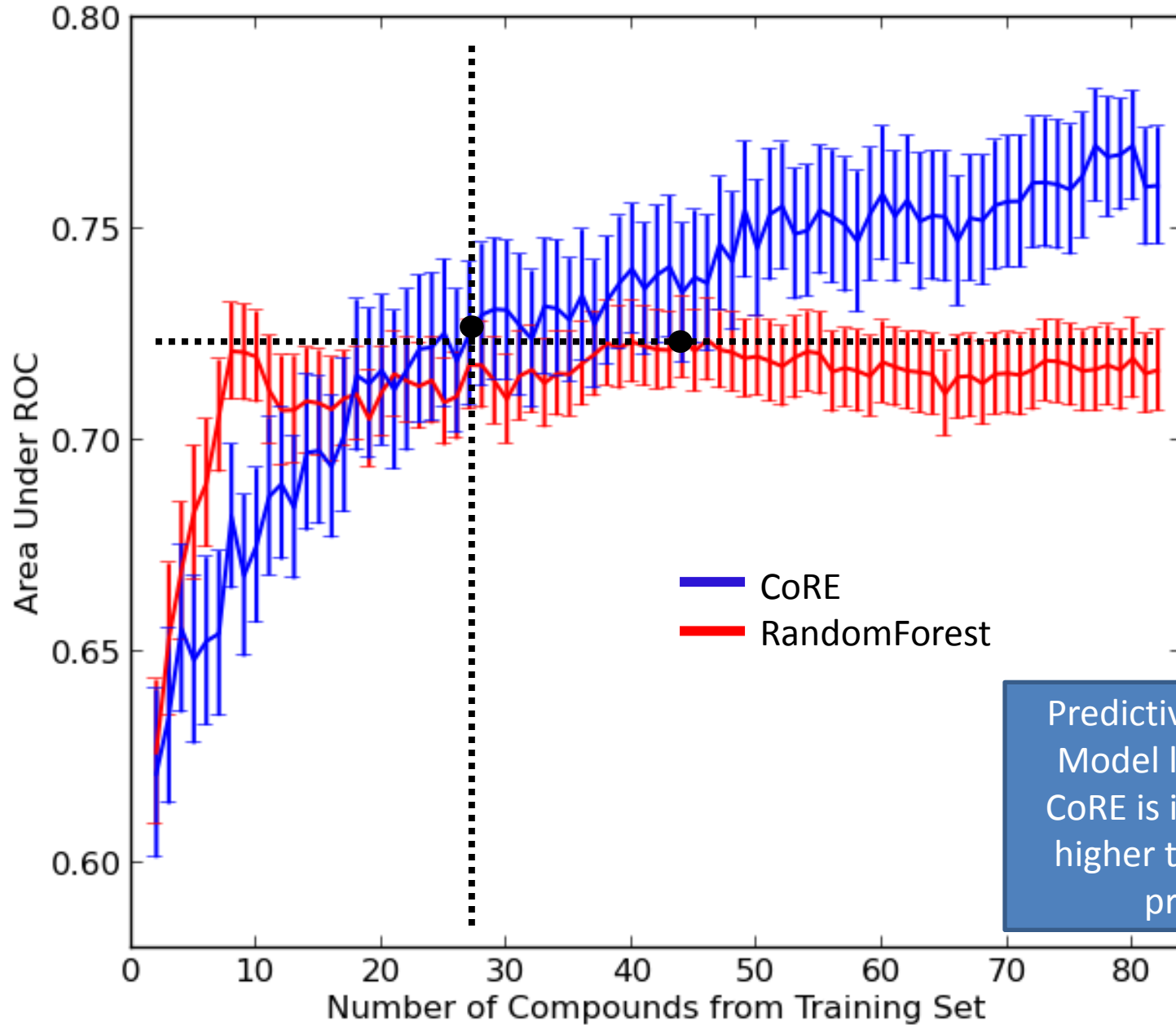




# Hepatotoxicity Simulation Results

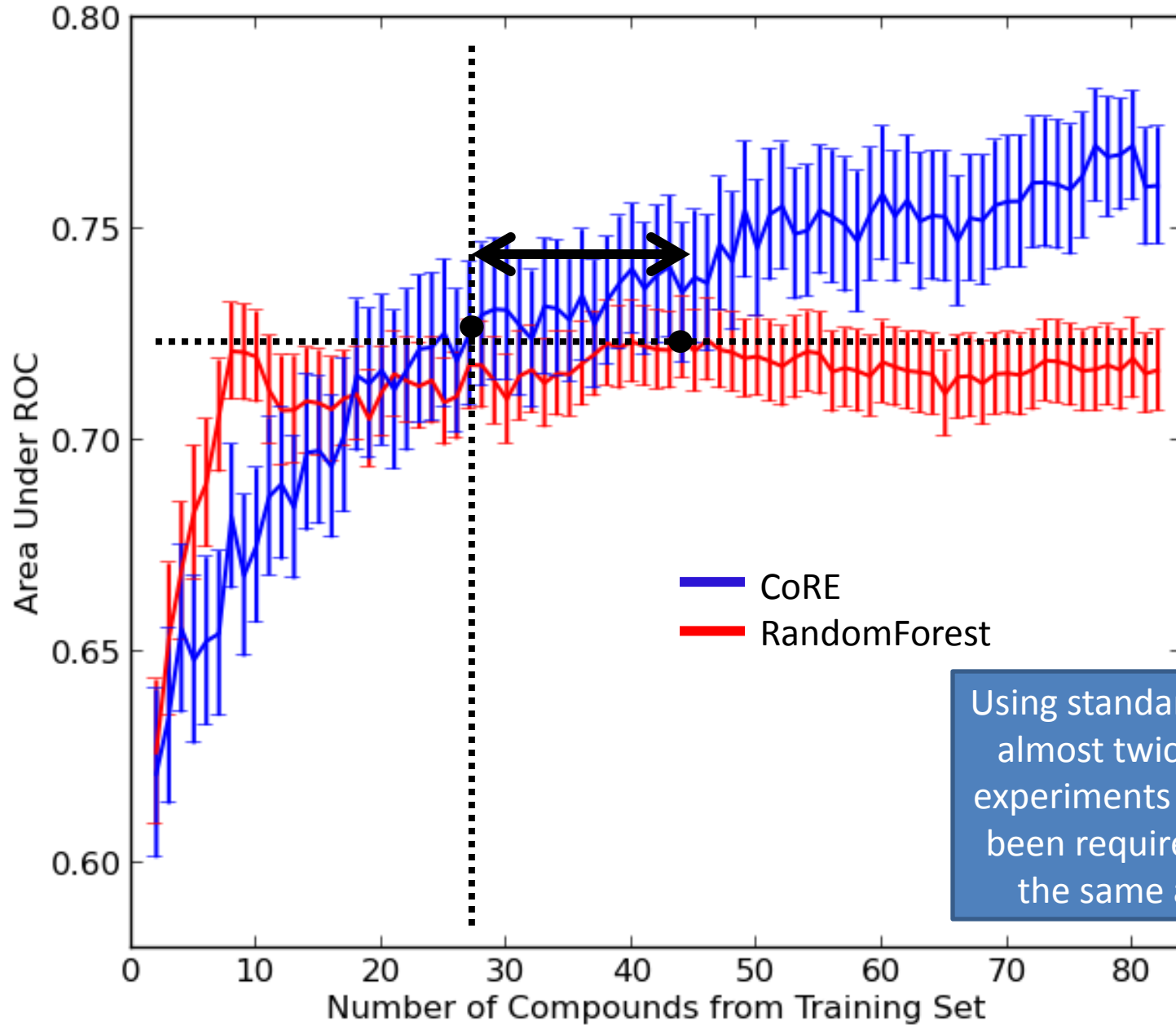


# Hepatotoxicity Simulation Results



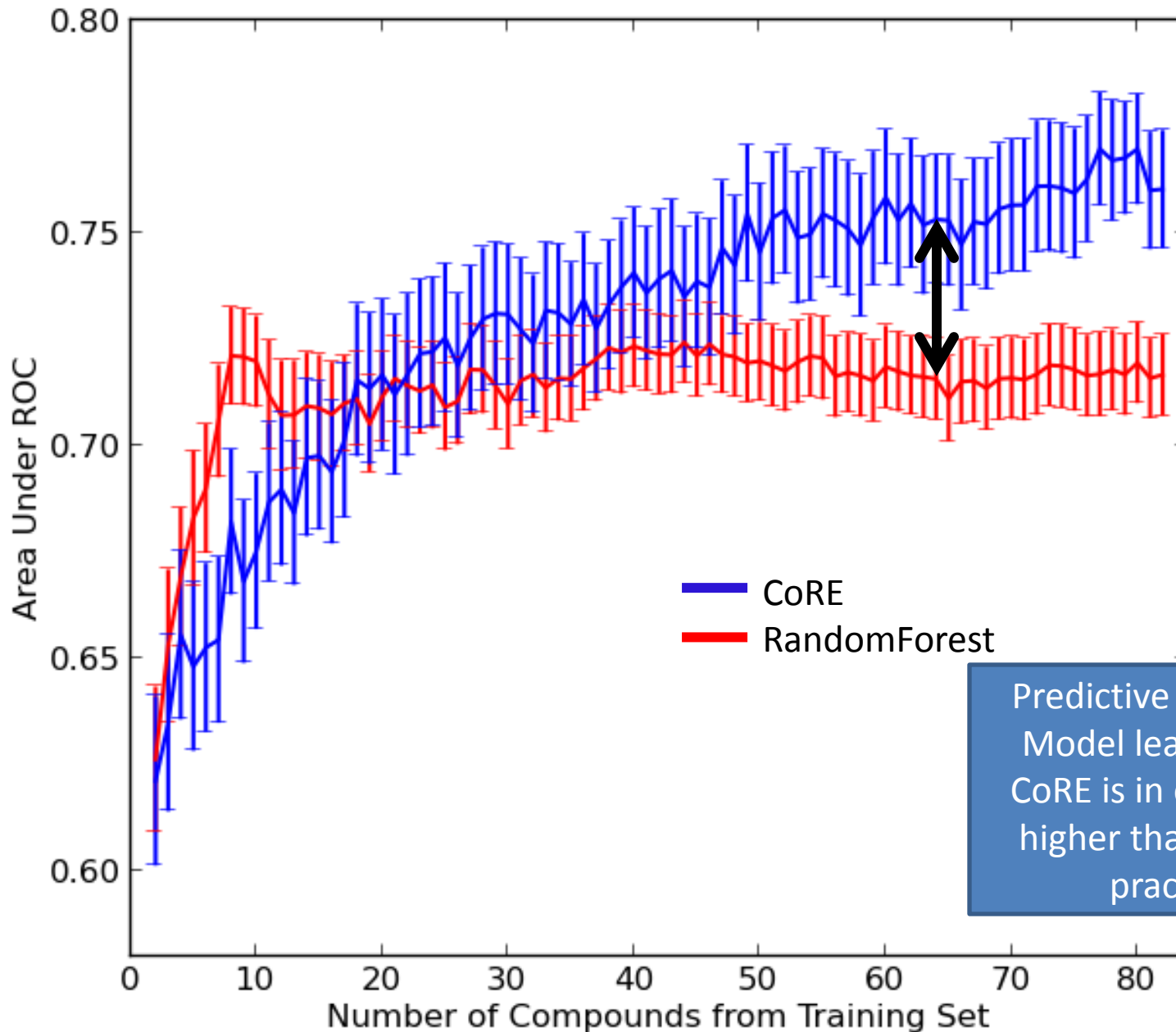
Predictive accuracy of Model learned using CoRE is in expectation higher than standard practices.

# Hepatotoxicity Simulation Results



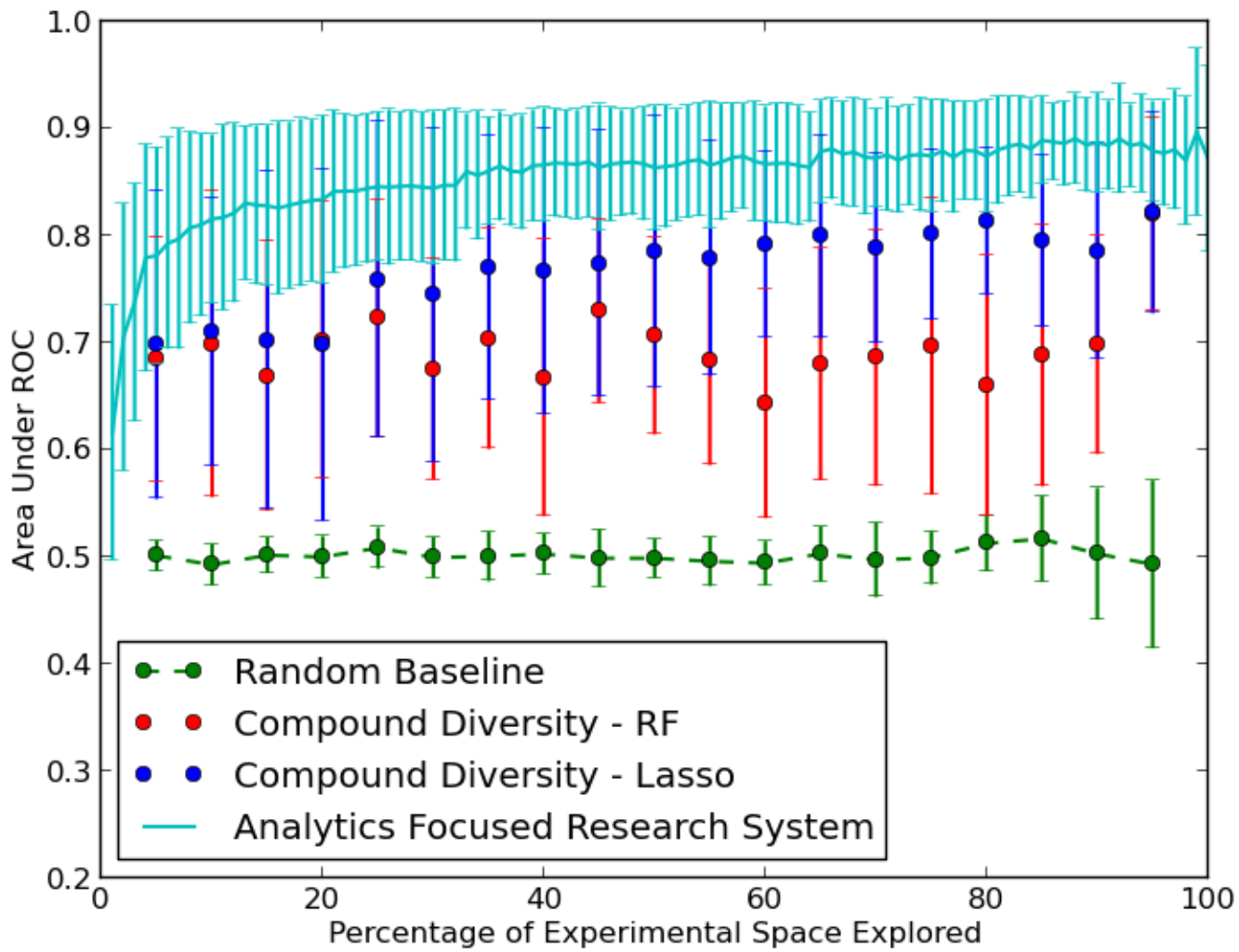
Using standard methods, almost twice as many experiments would have been required to reach the same accuracy.

# Hepatotoxicity Simulation Results

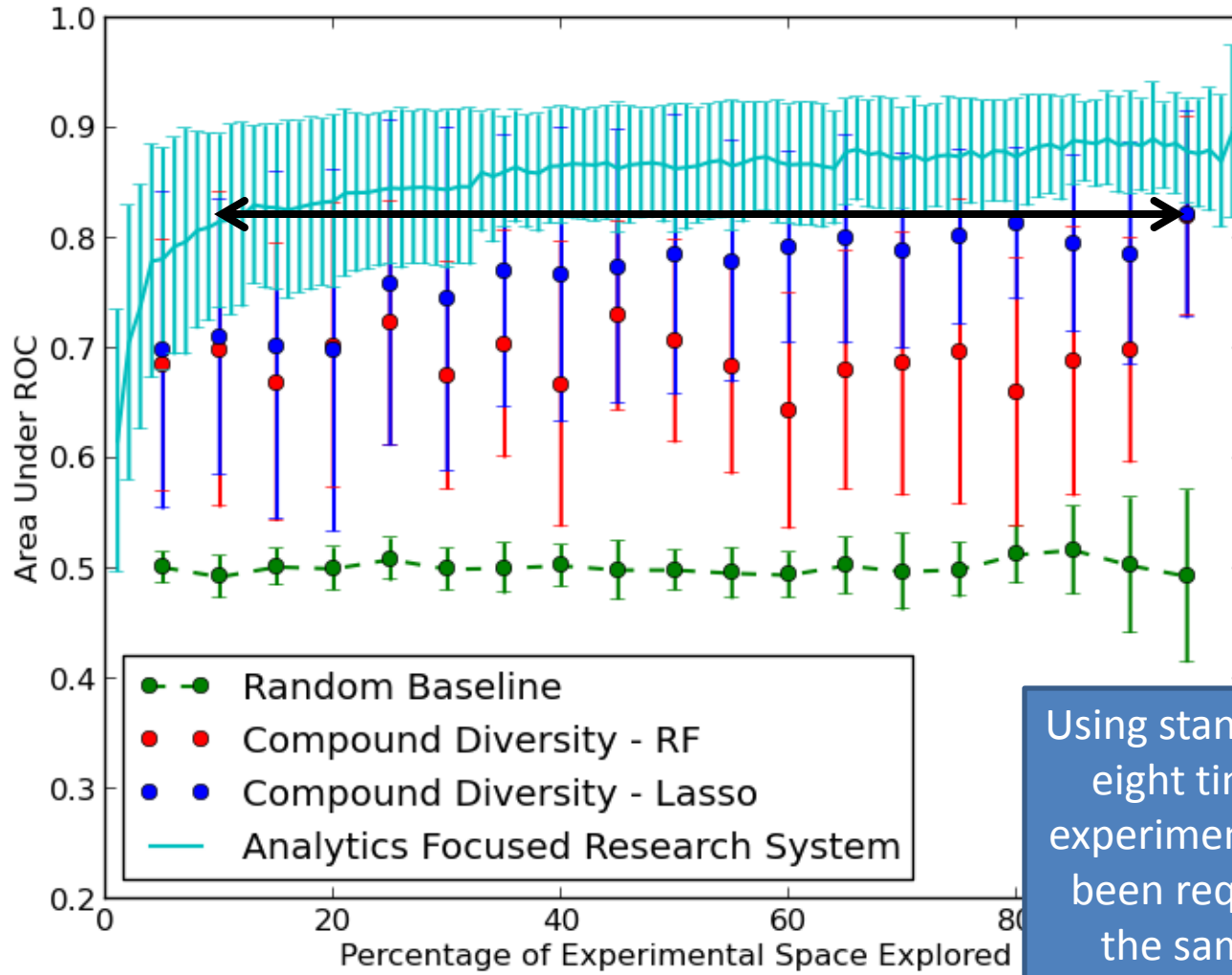


Predictive accuracy of Model learned using CoRE is in expectation higher than standard practices.

# ToxCast Simulation Results



# ToxCast Simulation Results



Using standard methods, eight times as many experiments would have been required to have the same accuracy.

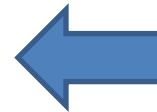
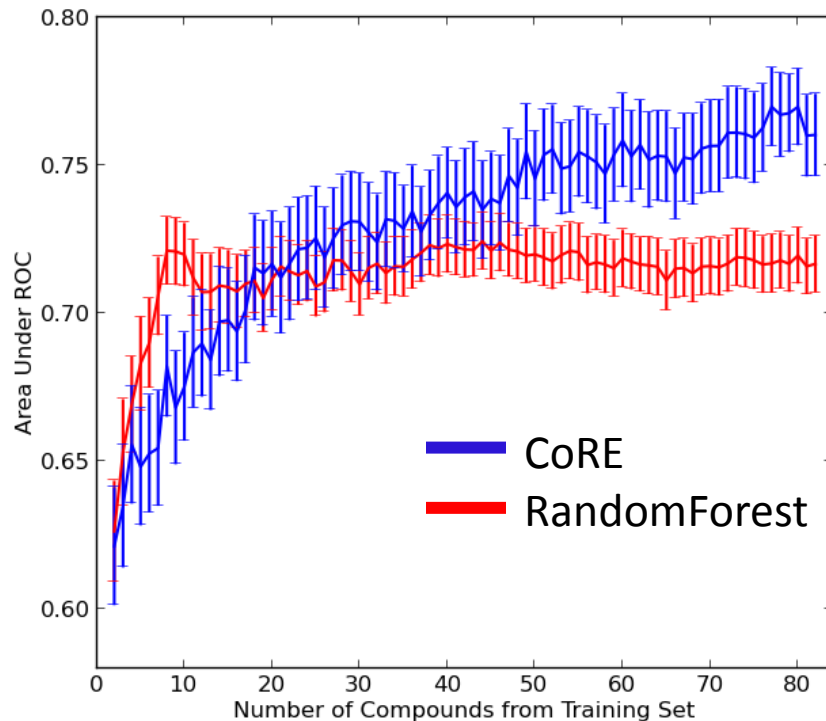
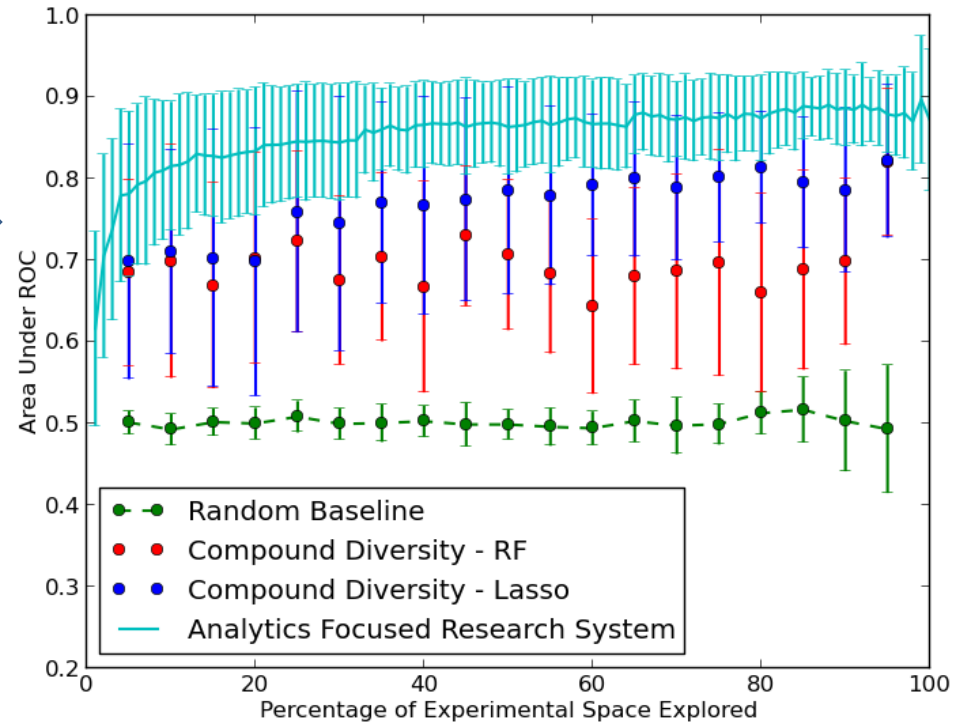
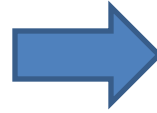
# Comparison with Similar Simulations

## ToxCast

Goal: Predict all unobserved results in ToxCast

1200 assays

309 compounds



## Current Study

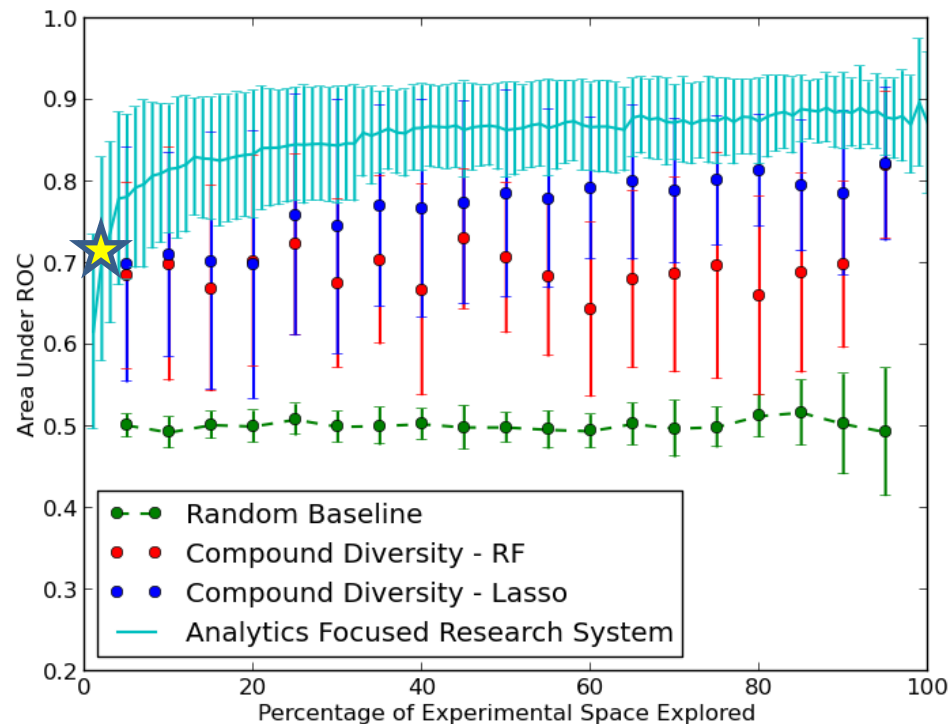
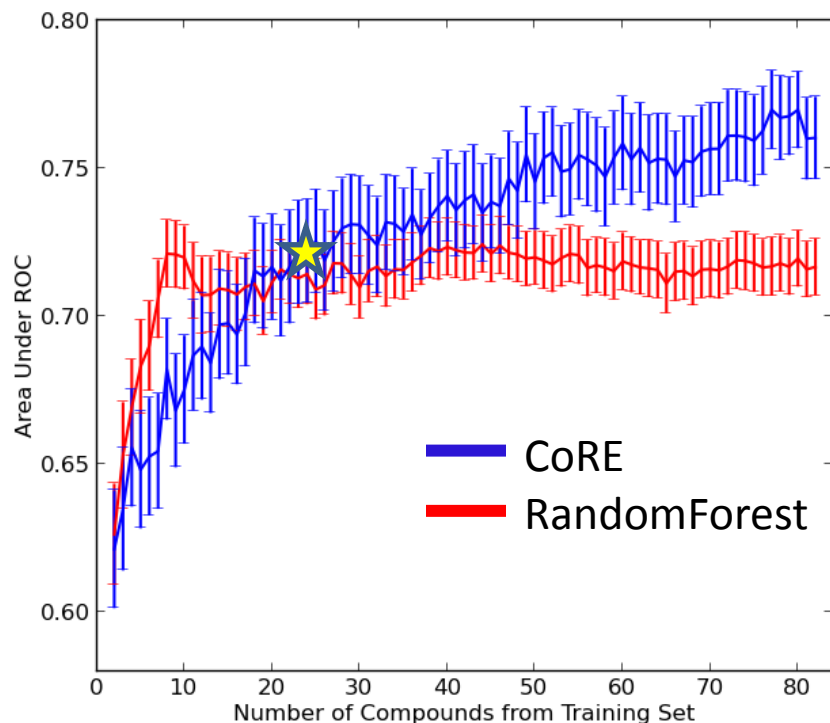
Goal: Predict toxicity

8 observations

105 compounds

# Comparison with Similar Simulations

What caused the difference in cross-over points?



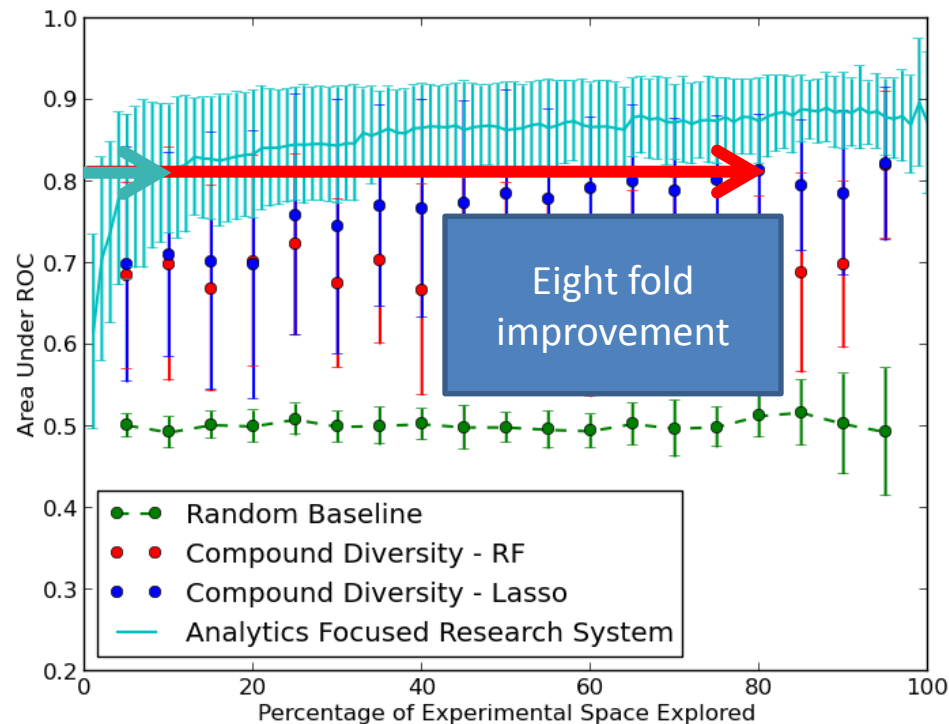
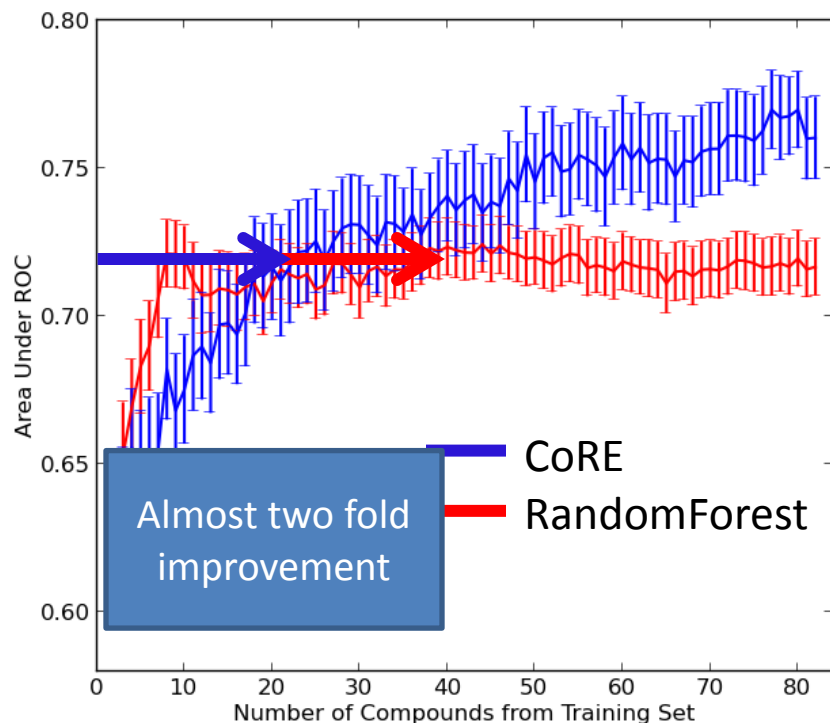
Potential Explanations:

- Quality of data
- Selection of compounds
- Dimensionality of experimental space
  - PubChem: 1200 measurements, 309 compounds
  - Current 8 measurements, 105 compounds



# Comparison with Similar Simulations

What cause the difference in the fraction of experiments required to reach maximum accuracy?



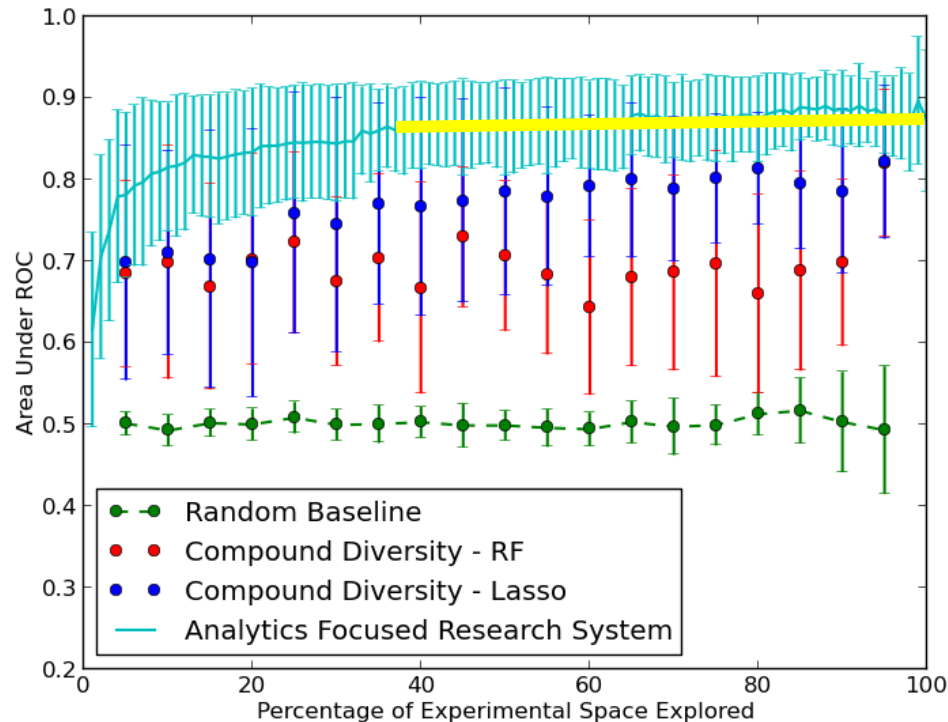
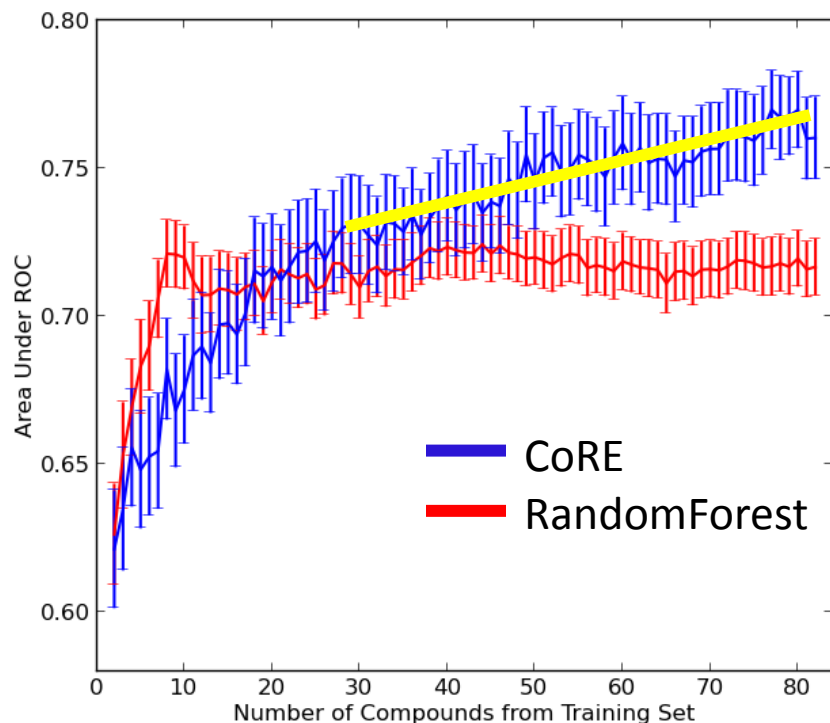
Potential Explanations:

- Quality of data
- Selection of compounds
- Dimensionality of experimental space
  - PubChem: 1200 measurements, 309 compounds
  - Current 8 measurements, 105 compounds

# Comparison with Similar Simulations

Why does accuracy in this study improve continuously?

Corollary: What is the upper limit to the achievable accuracy of toxicity prediction?



Potential Explanation:

- Dimensionality of experimental space
  - PubChem: 1200 measurements, 309 compounds
  - Current 8 measurements, 105 compounds

# Prospective Uses of Established CoRE Model

- For a single compound:
  - **More accurately predict the toxicity**
  - Prioritize toxicology assays to most effectively eliminate compounds (faster-to-fail)
  - Find multiple targets for drug discovery
- For multiple compounds:
  - **Prioritize compounds for further testing**
  - **Select experiments to effectively expand the predictions to new areas of chemical space**
- For multiple assays:
  - Prioritize assays based on human toxicity predictivity

# Next Studies?

- Prospective study
- Retrospective polypharmacology study
- Extend Retrospective toxicity study:
  - More compounds
  - More assays
  - More toxicities



# QUANTITATIVE MEDICINE

Transforming Drug Discovery™

SCOTT R. BODINE  
scott.bodine@qtmed.com

ROBERT F. MURPHY PHD  
robert.murphy@qtmed.com

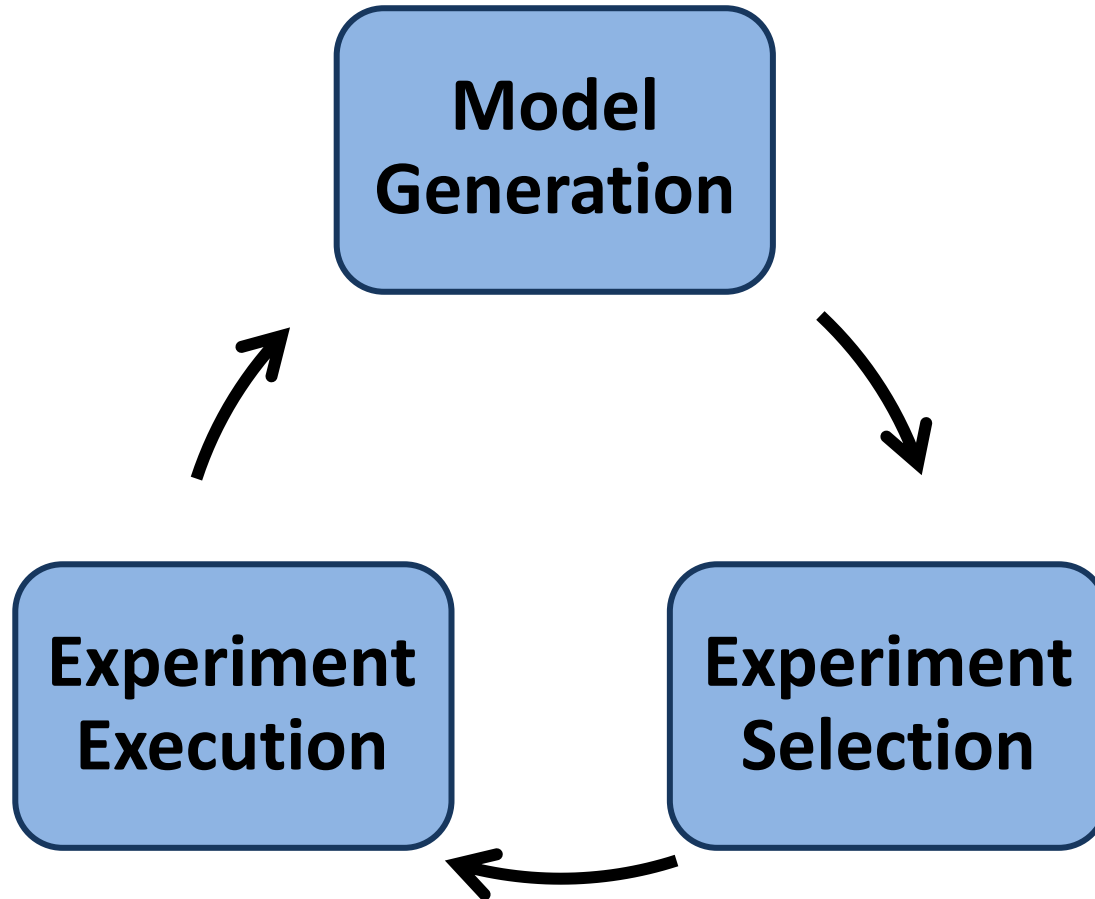
FOUNDERS

JOHN R. KING  
john.king@qtmed.com  
PRESIDENT & CEO

JOSHUA D. KANGAS PHD  
josh.kangas@qtmed.com  
CHIEF SCIENCE OFFICER

WWW.QTMED.COM

# Active Learning



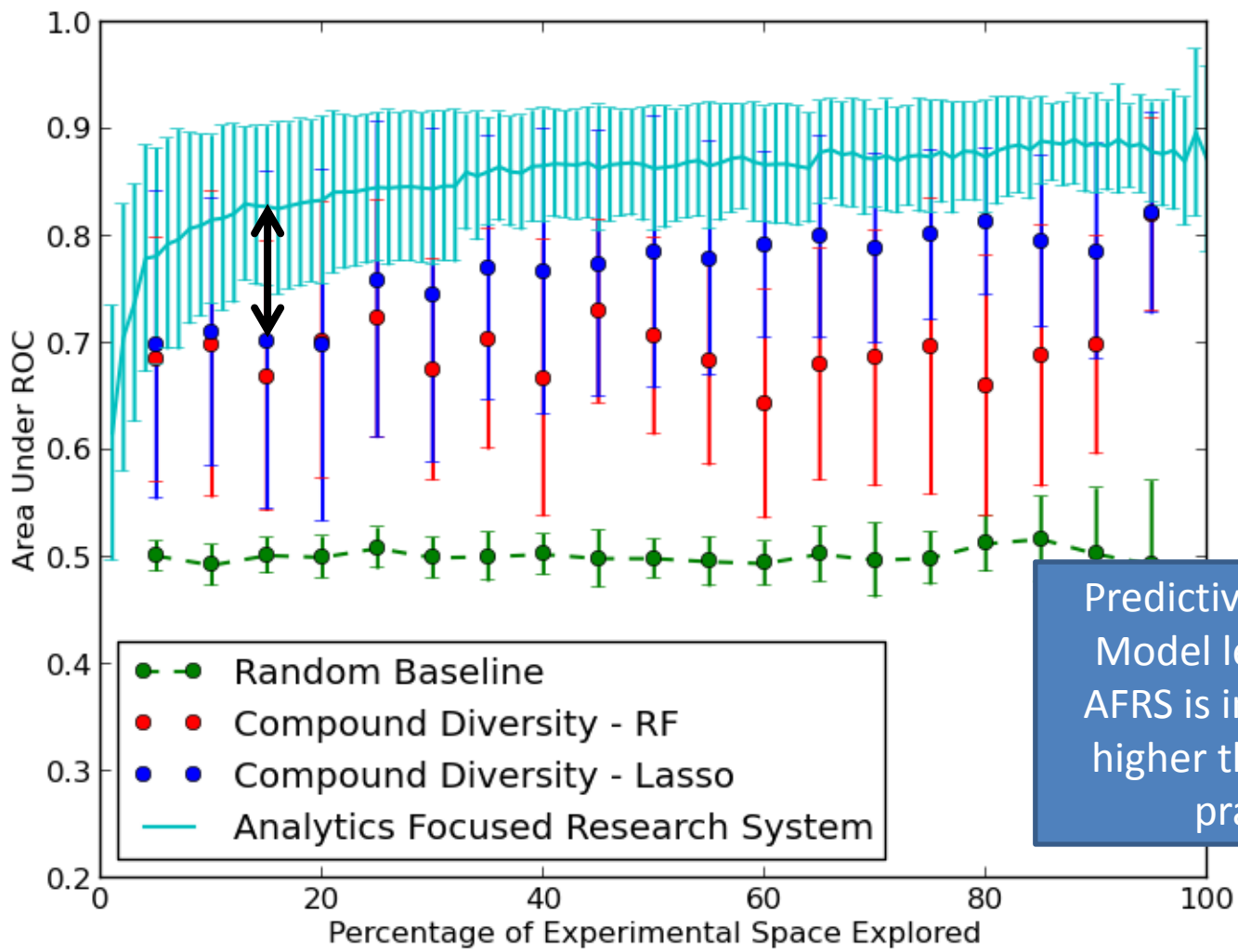
# AFRS Directed Study Types

## Retrospective

- Experiment Direction Simulated
- Show *potential* for improved **accuracy** and **efficiency**
- May discover new useful knowledge through AFRS analytics

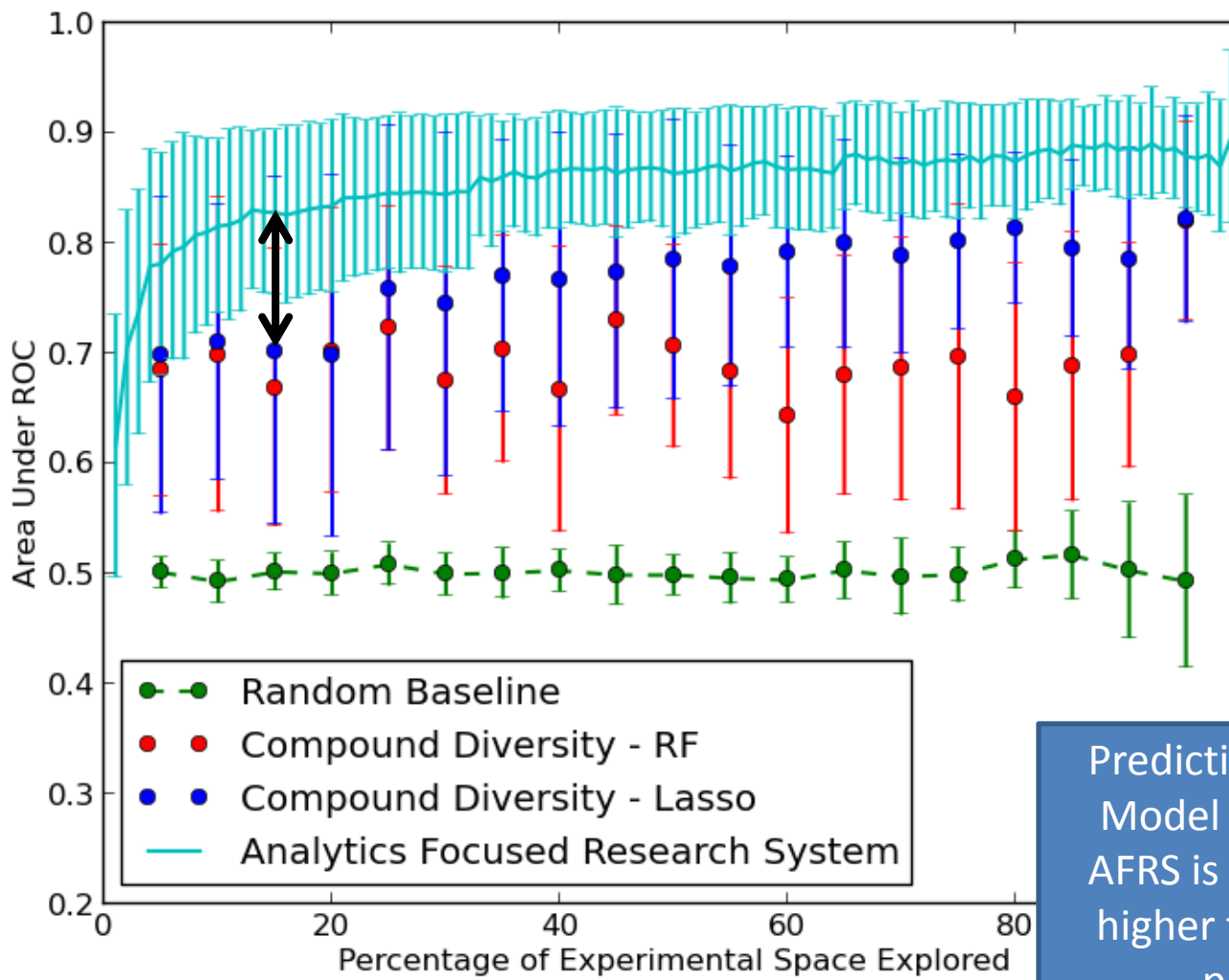
## Prospective

- Experiments Directed as Selected
- Yield *actual* improvements in accuracy and efficiency
- Likely to discover new useful knowledge through directed experimentation **and** AFRS analytics

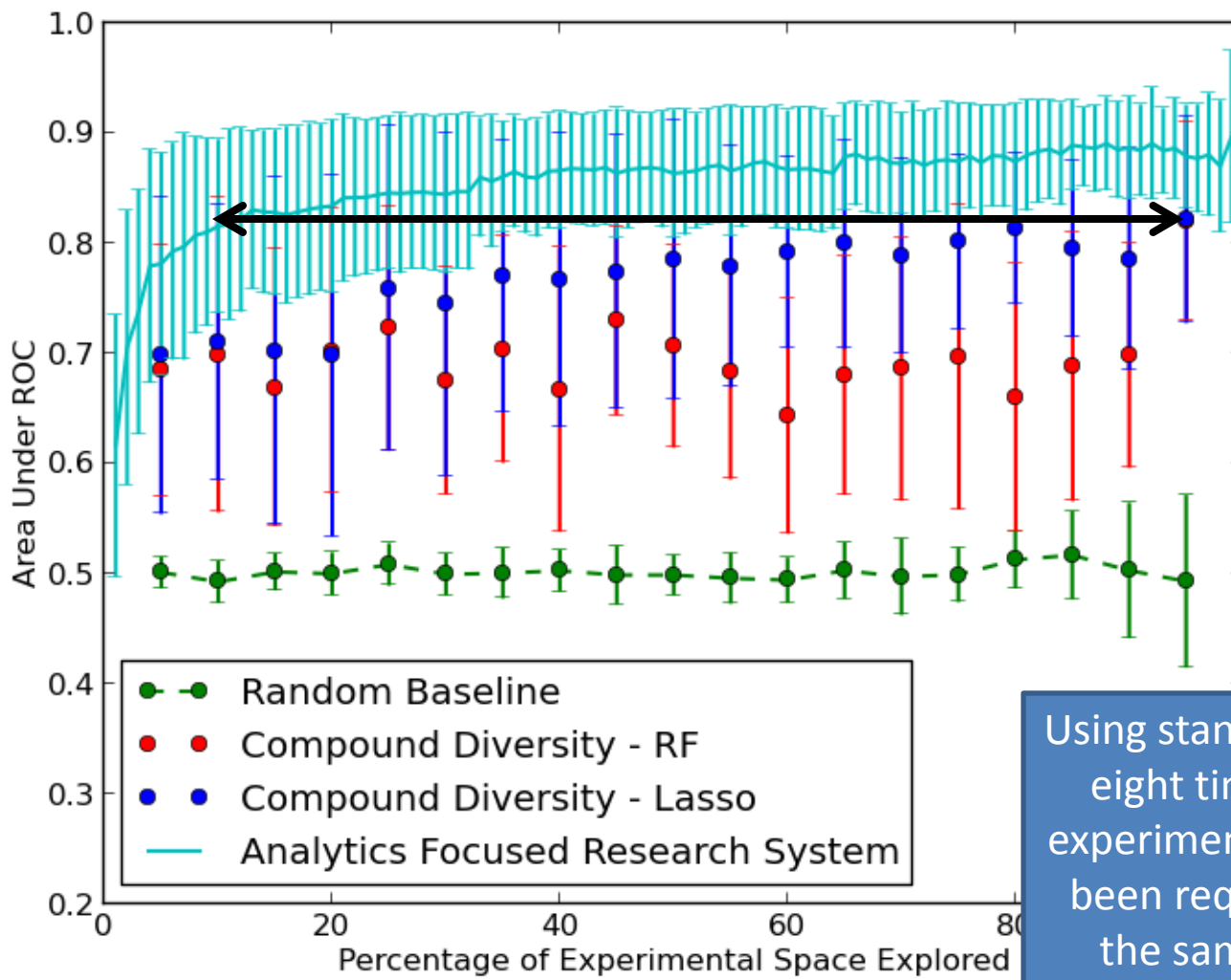


Predictive accuracy of Model learned using AFRS is in expectation higher than standard practices.





Predictive accuracy of Model learned using AFRS is in expectation higher than standard practices.



Using standard methods, eight times as many experiments would have been required to have the same accuracy.