Combining mechanistic modeling with machine learning to predict cardiotoxicity and streamline drug development

Eric Sobie, PhD
Professor, Pharmacological Sciences
Senior Associate Dean for Programmatic Development
eric.sobie@mssm.edu

Icahn School of Medicine at Mount Sinai

December 12, 2018

Cardiac electrophysiology QSP models

Inputs

- Electrical stimulation conditions (e.g. frequency)
- Temperature
- Ionic concentrations
- Drugs that may influence ion channels

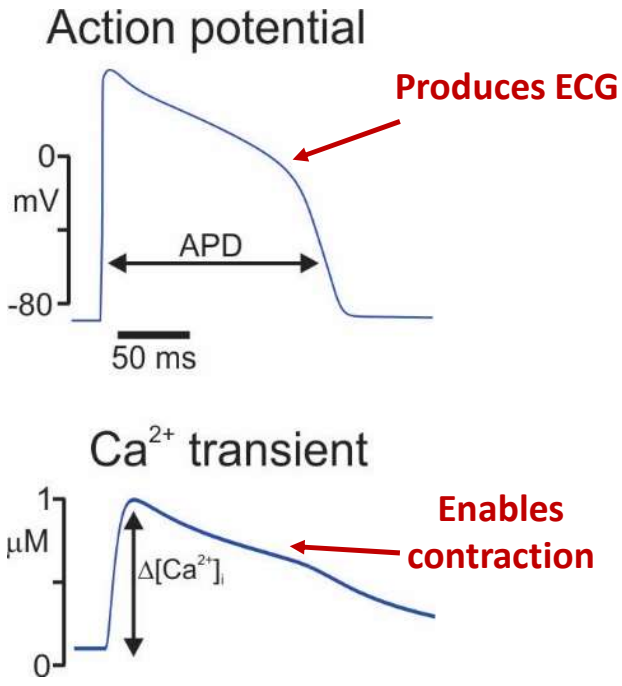
+

Mechanisms

- Current through ion channels, pumps, and transporters
- Ionic homeostasis

➡

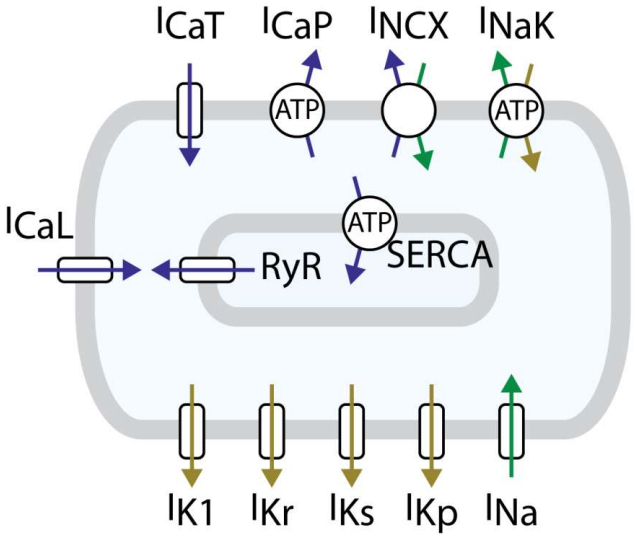
Outputs



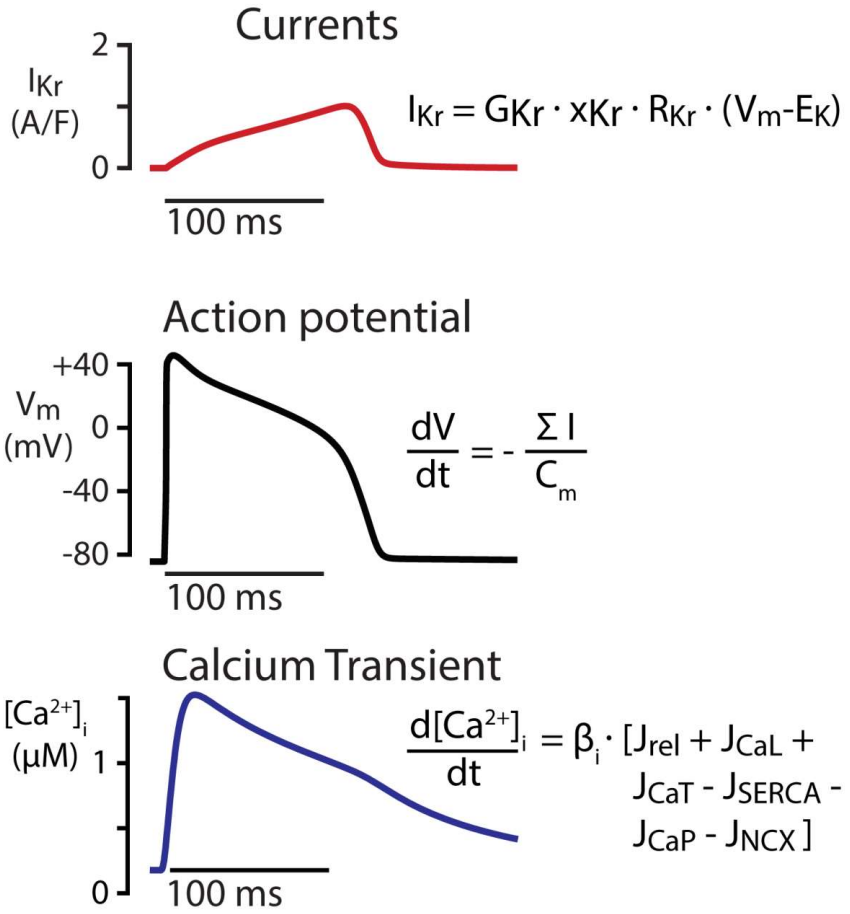
Mechanistic (QSP) cardiac myocyte models

Models simulate ionic currents, intracellular ionic homeostasis

Models have been developed over ~50 years of basic physiology research

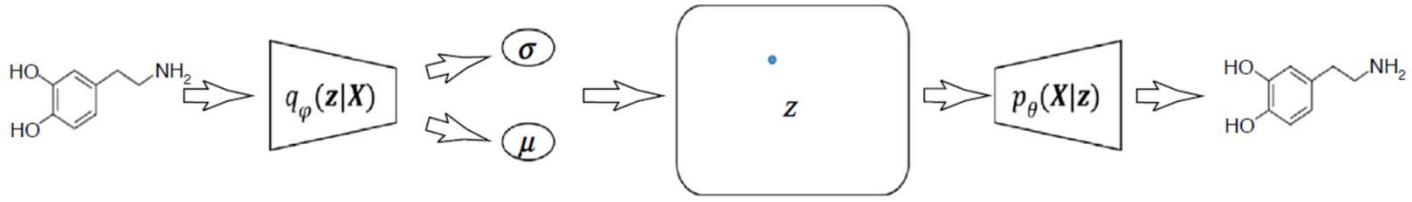


- 10-20 ion channels, pumps, and transporters
- 20-60 ordinary differential equations
- Drug effects simulated by reducing/enhancing activities



Machine Learning in drug development

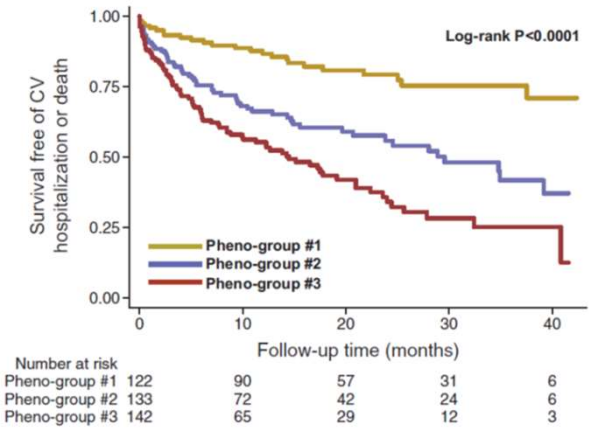
Analysis of drug chemical properties associated with outcomes



Chen et al *Drug Discovery Today* (2018) 23:1241.

Clustering of patients based on clinical data

Shah et al *Circulation* (2015) 131:269.



Is there any benefit in combining machine learning with mechanistic modeling?

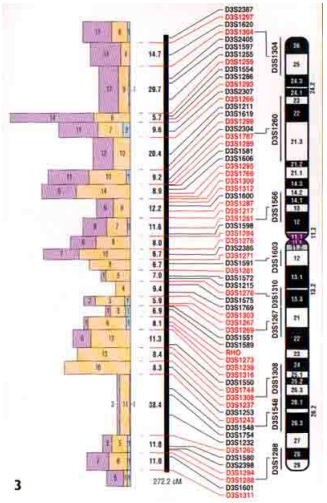
The Challenge of Systems Biology

The Omics Tradition

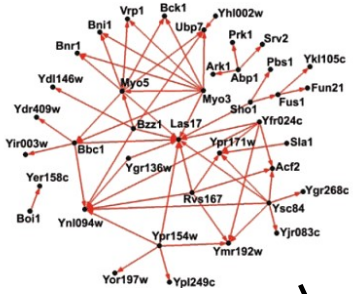
Large-scale
High-throughput

Descriptive
Qualitative

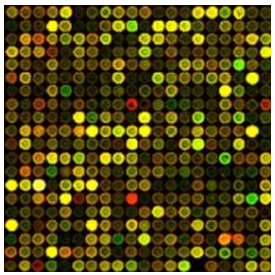
Human genome project



Networks



Microarrays



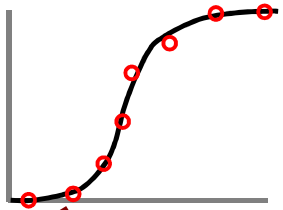
The Physiology Tradition

Small-scale
Low-throughput

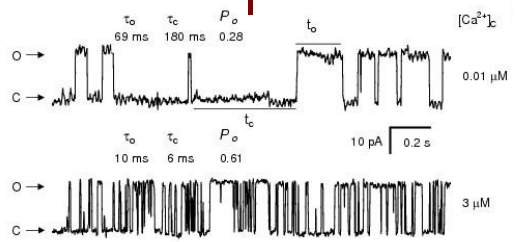
Mechanistic
Quantitative

Action potentials

Dose-response curves



Models



Patch clamp

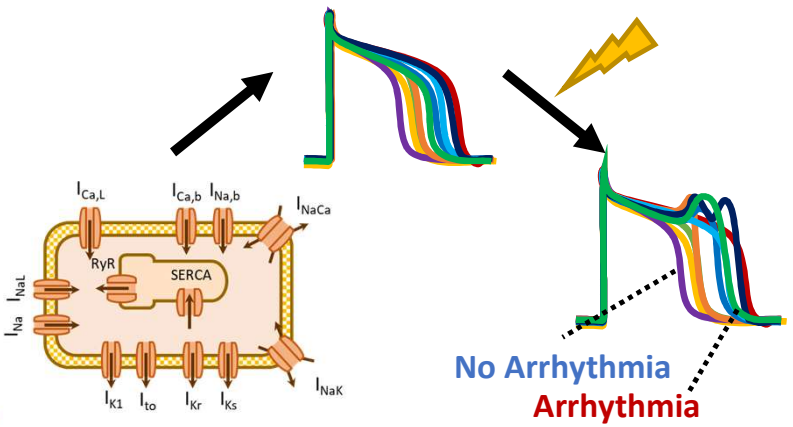
Systems Biology

Slide created December 2009

Strategy for integration of QSP and machine learning

QUANTITATIVE SYSTEMS PHARMACOLOGY

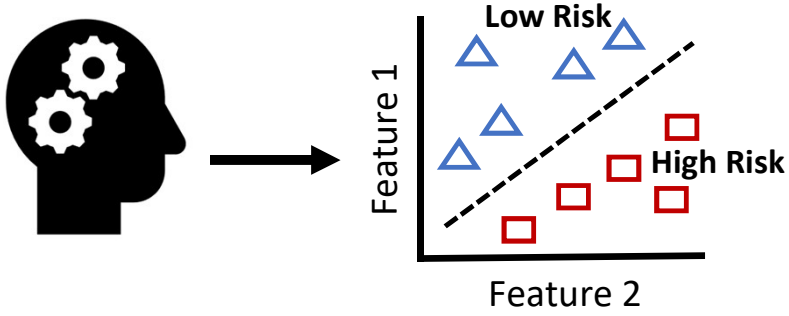
Build Population of Action Potentials



Extract features from population

MACHINE LEARNING ALGORITHMS

Determine best features to predict risk



1. **What mechanistic differences explain good or poor classification?**
2. **Which simulation protocols improve prediction?**

Questions we can address with machine learning

Can we streamline and improve testing of drugs for potential proarrhythmic effects?

Can we correct for limitations of experimental models?

Can we distinguish between susceptible and resistant patient groups?

Outline

Computational prediction of proarrhythmia

Megan Cummins Lancaster, MD/PhD
Vanderbilt University

Cummins Lancaster & Sobie, *Clinical Pharmacology & Therapeutics* (2016) 100:371-379.

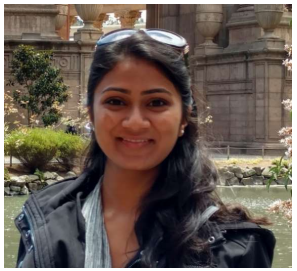
Predicting drug effects across cell types

Jingqi Gong
PhD candidate

Gong & Sobie *npj Systems Biology & Applications* (2018) 4:11.

Identification of susceptible sub-populations

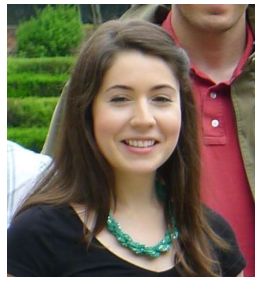
Varshneya, Mei, & Sobie, unpublished work in progress



Meera Varshneya
PhD candidate

Outline

Computational prediction of proarrhythmia

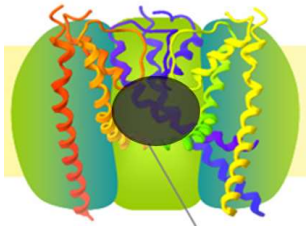


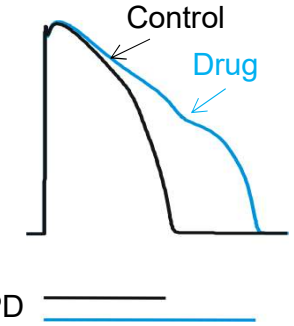
Megan Cummins Lancaster, MD/PhD
Vanderbilt University

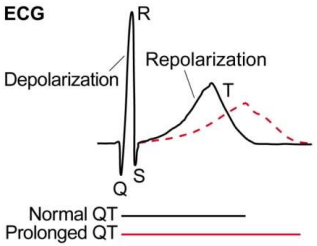
Cummins Lancaster & Sobie, *Clinical Pharmacology & Therapeutics* (2016) 100:371-379.

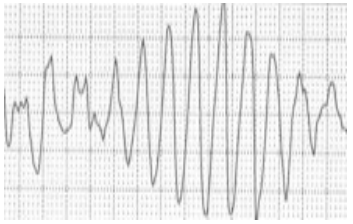
Drug-induced Torsades de Pointes

Blockade of cardiac ion channels, especially hERG (I_{Kr}) can prolong action potentials and cause arrhythmias.









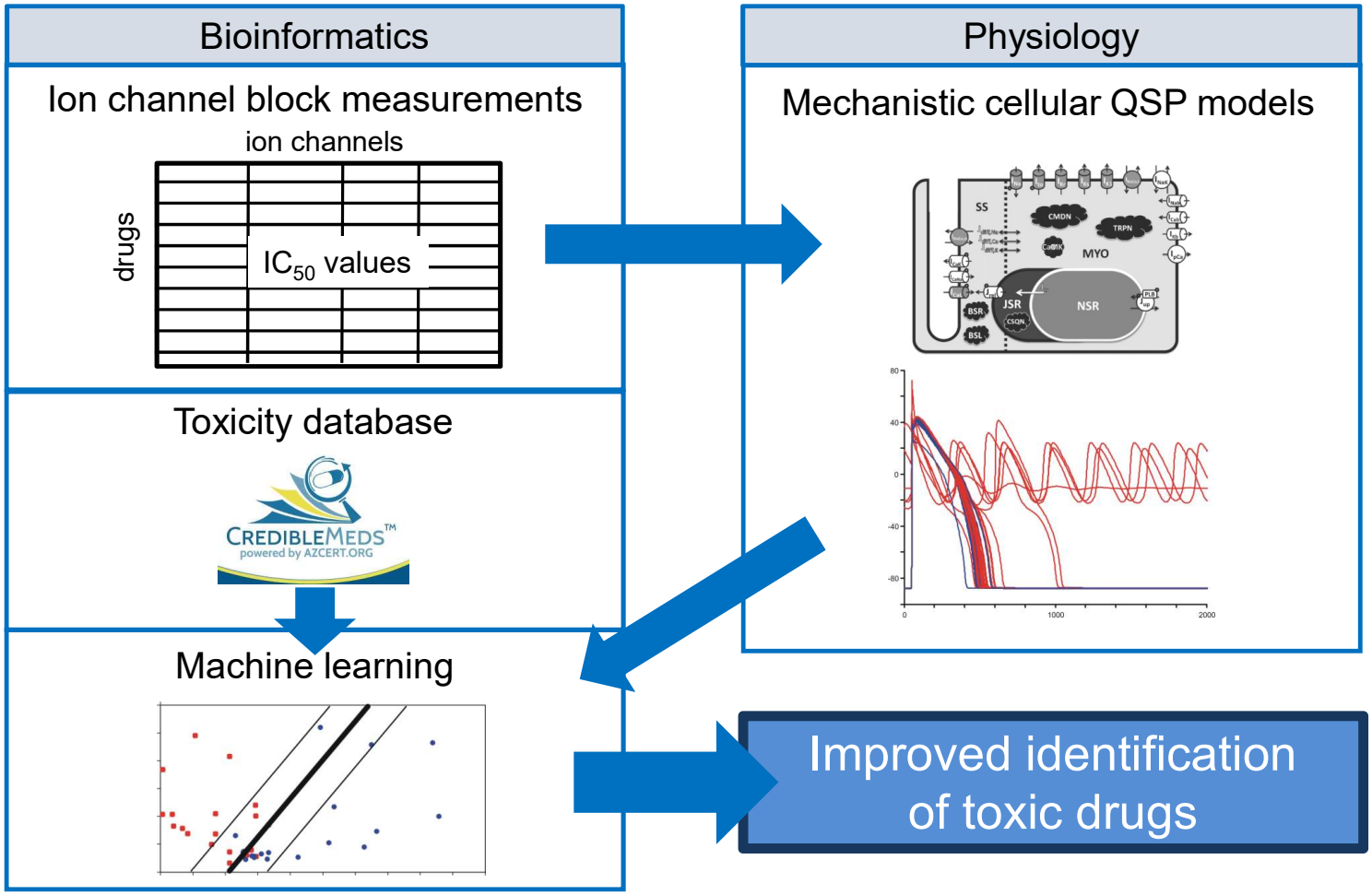
hERG block → Action potential lengthening → QT prolongation → Torsades de Pointes

Complicating factors:

- **Drugs block multiple ion channels**
- **Arrhythmias can arise without substantial action potential prolongation**
- **Both cellular and tissue effects can contribute**

hERG image: Grilo et al., (2010) *Front in Pharm* 1:137

QSP modeling to improve Torsades prediction



Develop algorithm based on multi-channel block

Model drug interaction with multiple ion channels

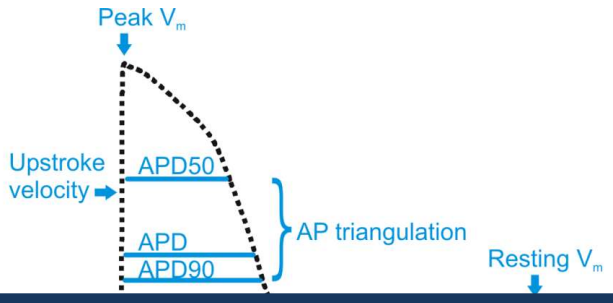
Drug	hERG IC50	Nav1.5 IC50	Cav1.2 IC50	TdP Risk
Ibutilide	0.018	42.5	62.5	1
Quinidine	0.72	14.6	6.4	1
Verapamil	0.25	32.5	0.2	0
Flecainide	1.5	6.2	27.1	1
Nifedipine	44	88.5	0.012	0
Thioridazine	0.5	1.4	3.5	1
….				

67 drugs: blocking potency of 3 ion channels and clinical TdP risk

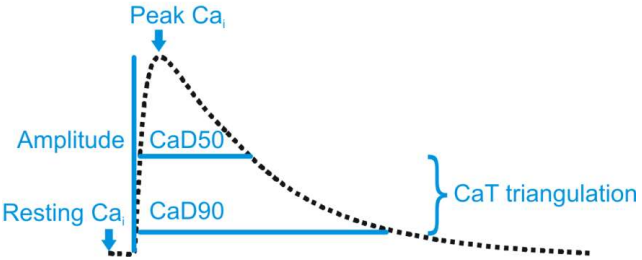
Sources:
Mirams GR et al.(2011). *Cardiovasc Res* 91: 53-61.
Kramer J et al.(2013). *Sci Rep* 3: 2100.
Champeroux P et al. (2011). *J Pharm and Tox Meth* 63: 269-278.
CredibleMeds QT Drugs List. Oro Valley, AZ: AZCERT, Inc.

Simulations produce a large set of pseudo-data

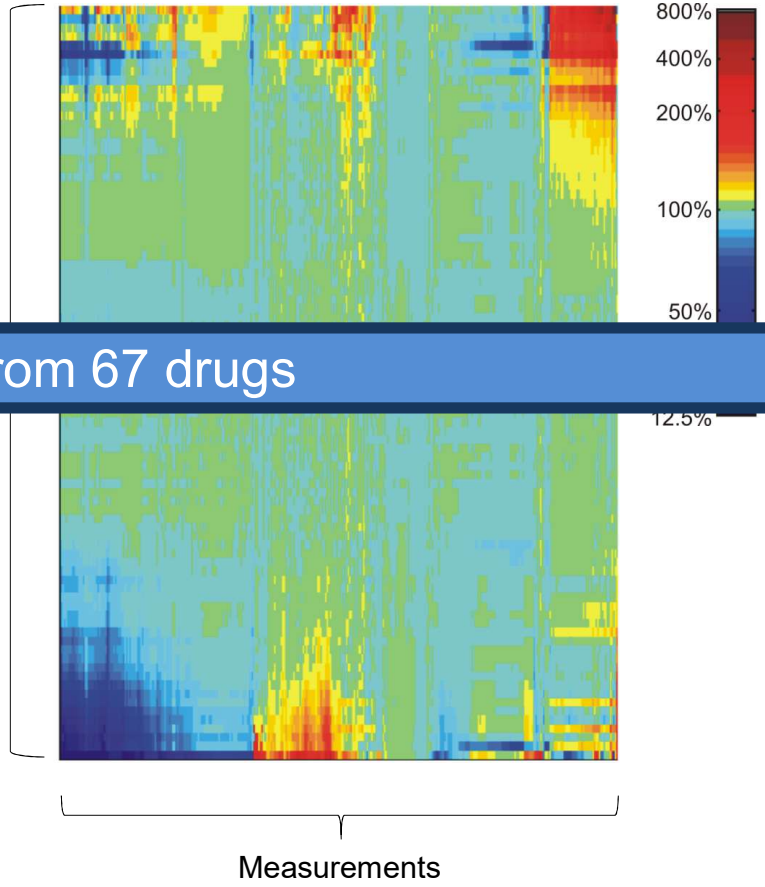
Action potential:



Ca^{2+} transient:



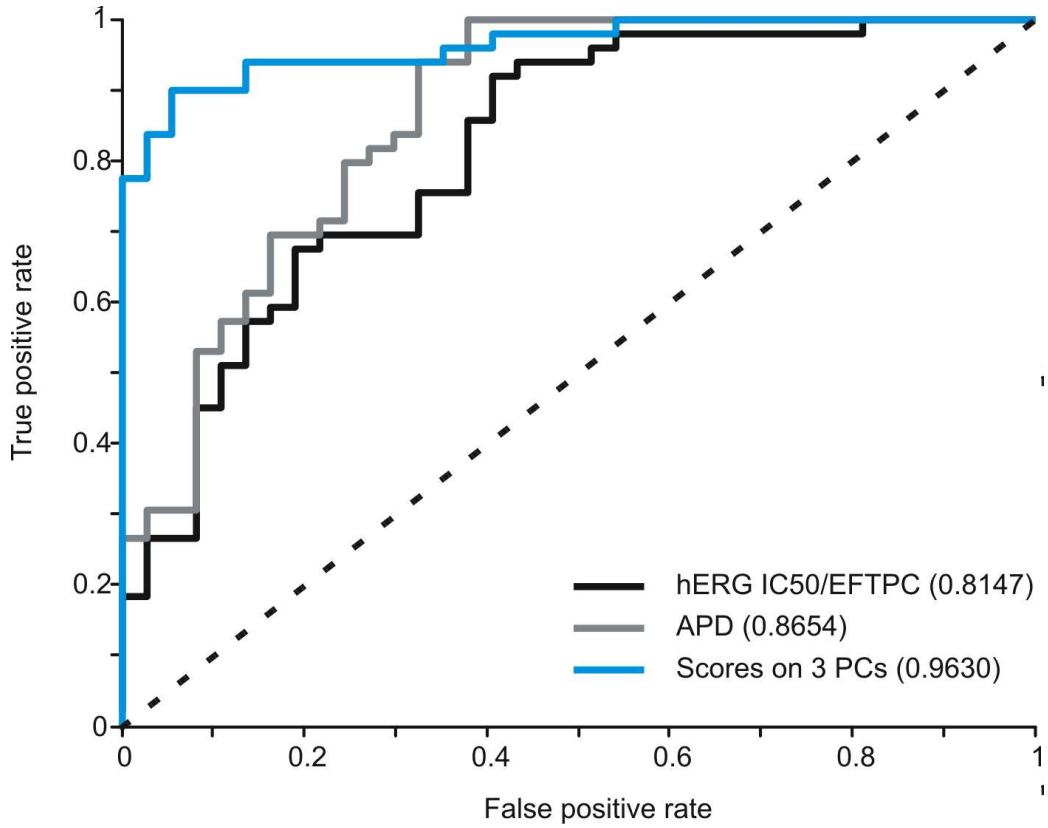
Also AP rate dependence and transmural dispersion



331 metrics from 67 drugs

Cummins Lancaster & Sobie, *Clinical Pharmacology & Therapeutics* (2016) 100:371-379.

Machine Learning produces an improved classifier



auROC = 0.8147

auROC = 0.8654

auROC = 0.9630

Superior prediction of arrhythmia risk for real drugs.

QSP modeling adds value besides prediction

Dose dependence: **The simulations predict that some drugs only reveal arrhythmia risk at high plasma concentrations.**

Precision predictions: **Analysis of the simulation results provides insight into *which individuals* are at greatest risk of drug-induced Torsades.**

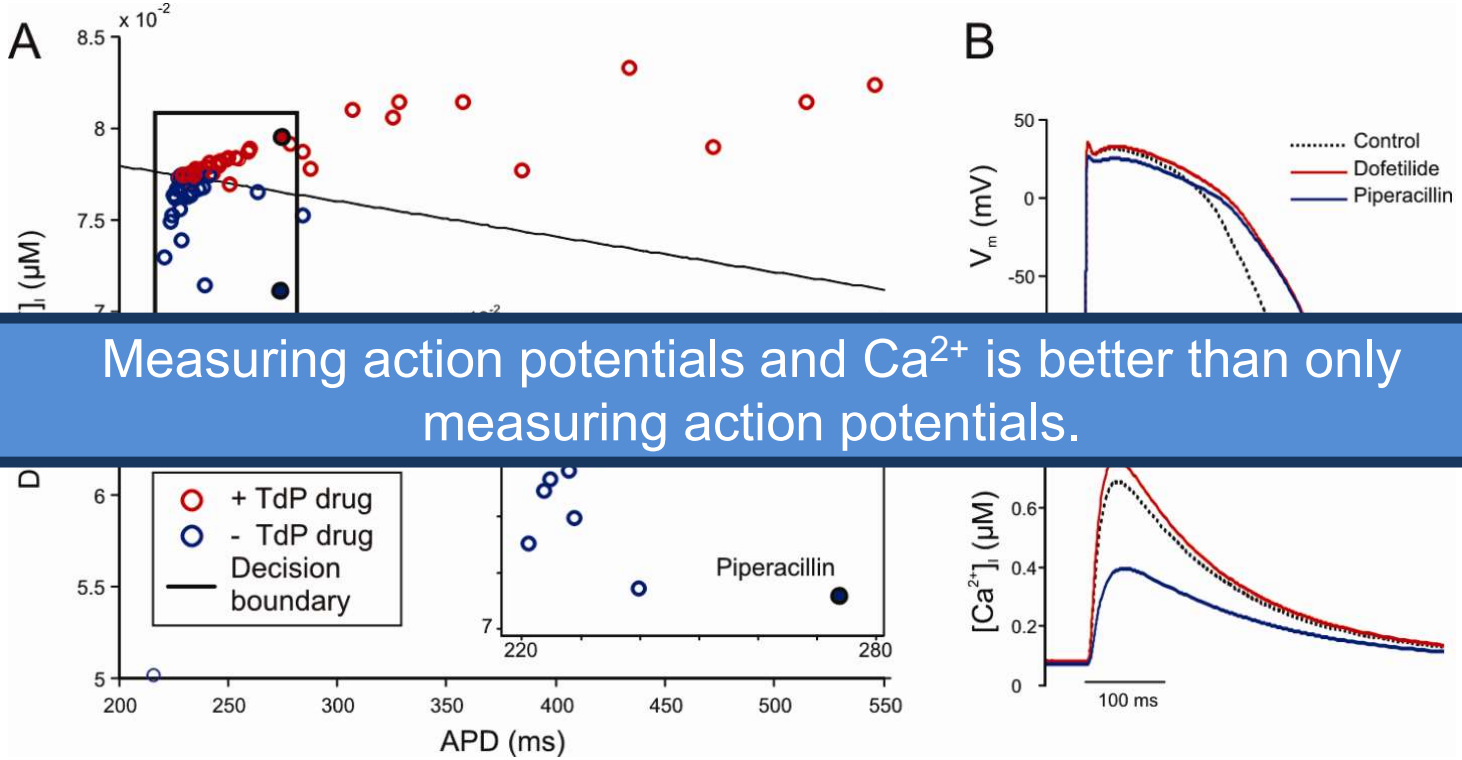
Experimental prioritization: **The simulations predict *which ion channels* should be assessed, and *which assays* should be conducted during *in vitro* physiology experiments.**

Mapping classification to physiology

Can we discriminate drugs based on physiological metrics?

Systematically determine most informative pairs of uncorrelated metrics

Top pair: diastolic $[Ca^{2+}]_i$ and APD_{50} in O'Hara epicardial model at 1 Hz



Measuring action potentials and Ca^{2+} is better than only measuring action potentials.

Cummins Lancaster & Sobie, *Clinical Pharmacology & Therapeutics* (2016) 100:371-379.

Drug classification study – key takeaway

The simulations do not only assist in classifying drugs (+TdP versus -TdP) -- they also indicate which assays are most helpful, and which ion channels contribute to toxicity.

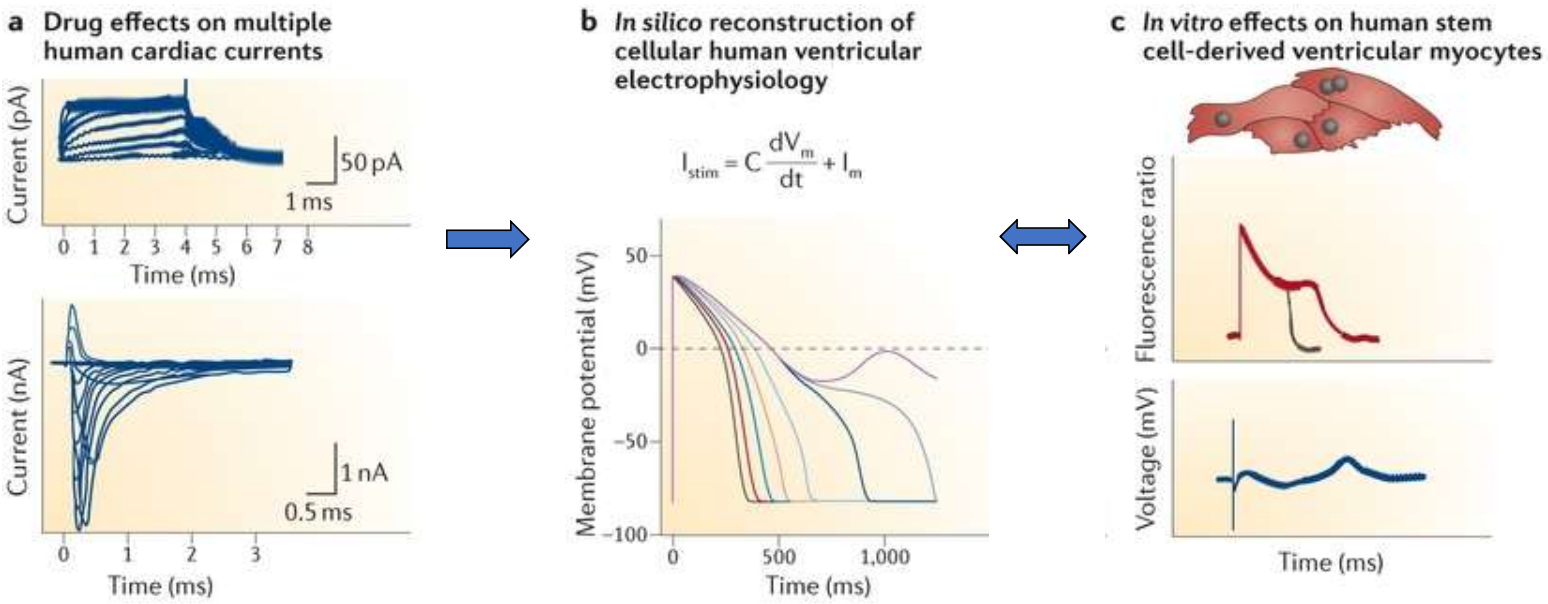
The role of modeling in Torsades prediction

CiPA = Comprehensive *in vitro* Proarrhythmia Assay

Ion channel block

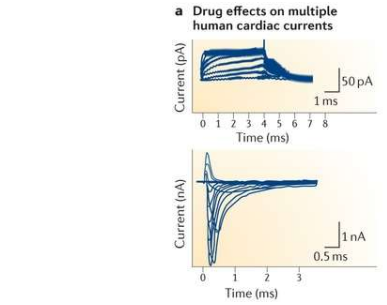
Simulations

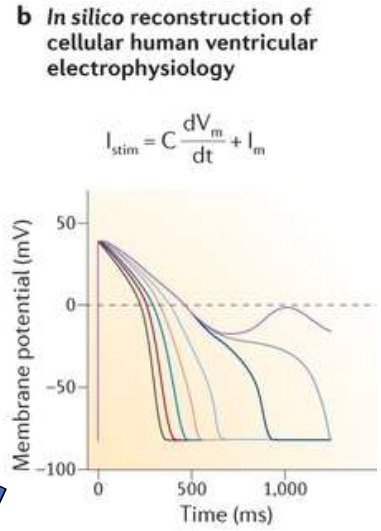
Cellular physiology

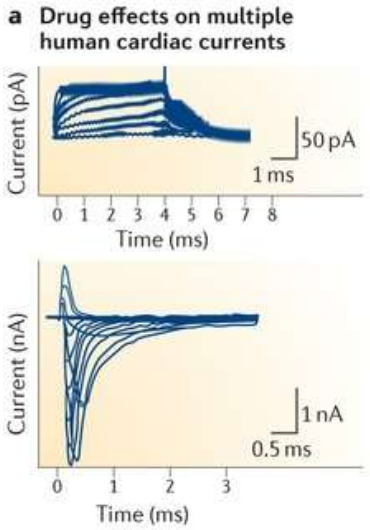


Gintant et al, *Nat. Rev. Drug Discovery* (2016) 15:457-471.

Proposal: an enhanced role for modeling in CiPA

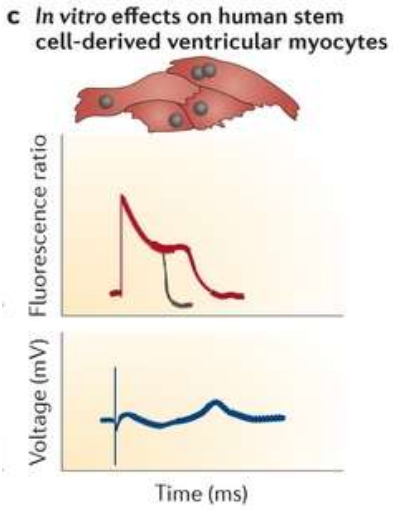






which channels?

which assays?

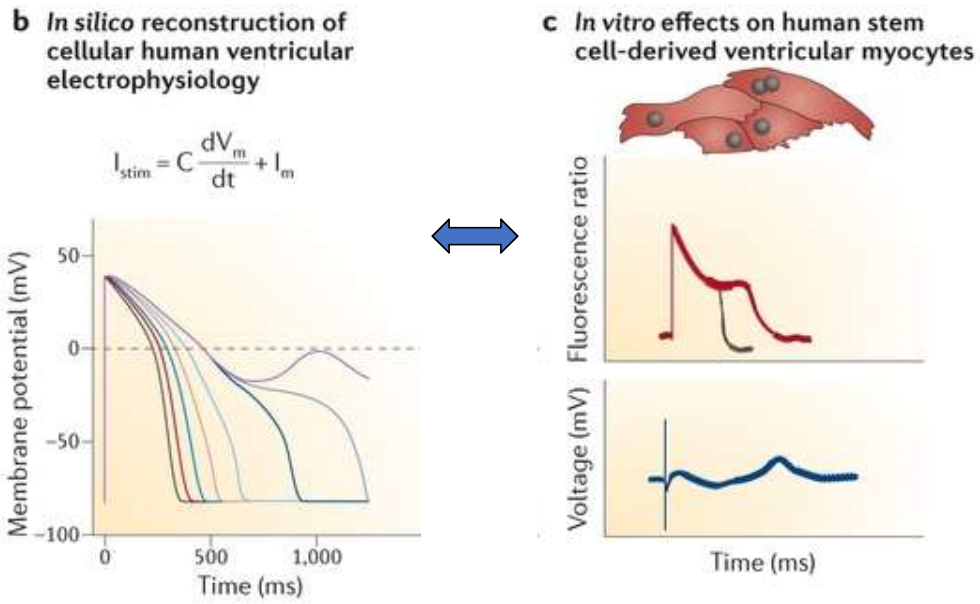


The role of modeling in Torsades prediction

CiPA = Comprehensive *in vitro* Proarrhythmia Assay

Simulations

Cellular physiology



Gintant et al, *Nat. Rev. Drug Discovery* (2016) 15:457-471.

Outline

Computational prediction of proarrhythmia

Megan Cummins Lancaster, MD/PhD
Vanderbilt University

Cummins Lancaster & Sobie, *Clinical Pharmacology & Therapeutics* (2016) 100:371-379.

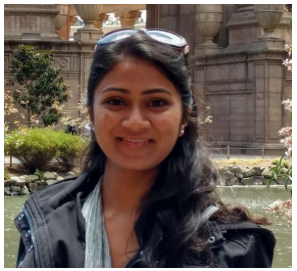
Predicting drug effects across cell types

Jingqi Gong
PhD candidate

Gong & Sobie *npj Systems Biology & Applications* (2018) 4:11.

Identification of susceptible sub-populations

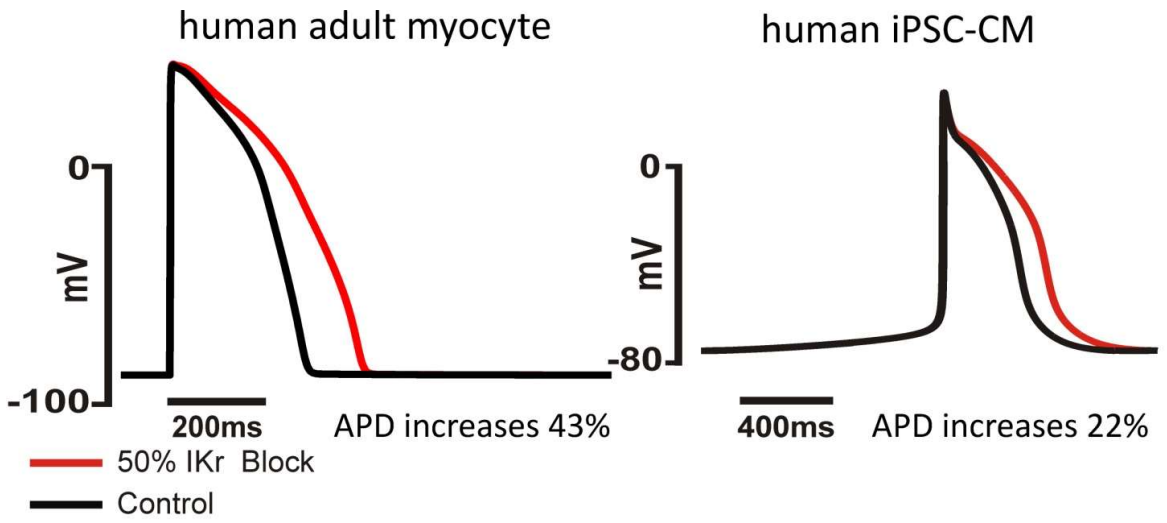
Varshneya, Mei, & Sobie, unpublished work in progress



Meera Varshneya
PhD candidate

Quantitative differences in simulated drug responses

I_{Kr} = rapid delayed rectifier K^+ current (hERG)



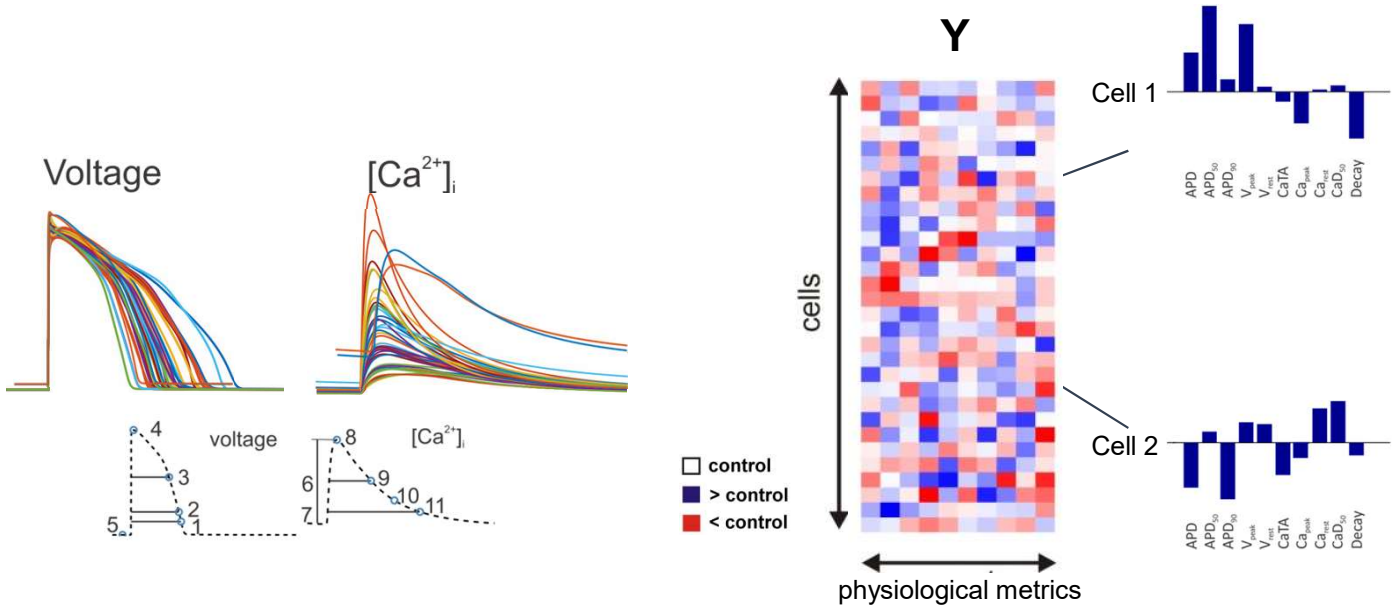
O'Hara et al. *PLOS Comp. Bio.* (2011) 7:e1002061. Paci et al. A*nn BME (*2013) 41:2334-2348.

Hypothesis:

Drug effects can be ***quantitatively translated*** from one cell type to another through a model that combines population-based mechanistic simulations and machine learning.

Population-based simulations to incorporate variability

Heterogeneity is generated by randomly varying the parameters that control ion channel abundances



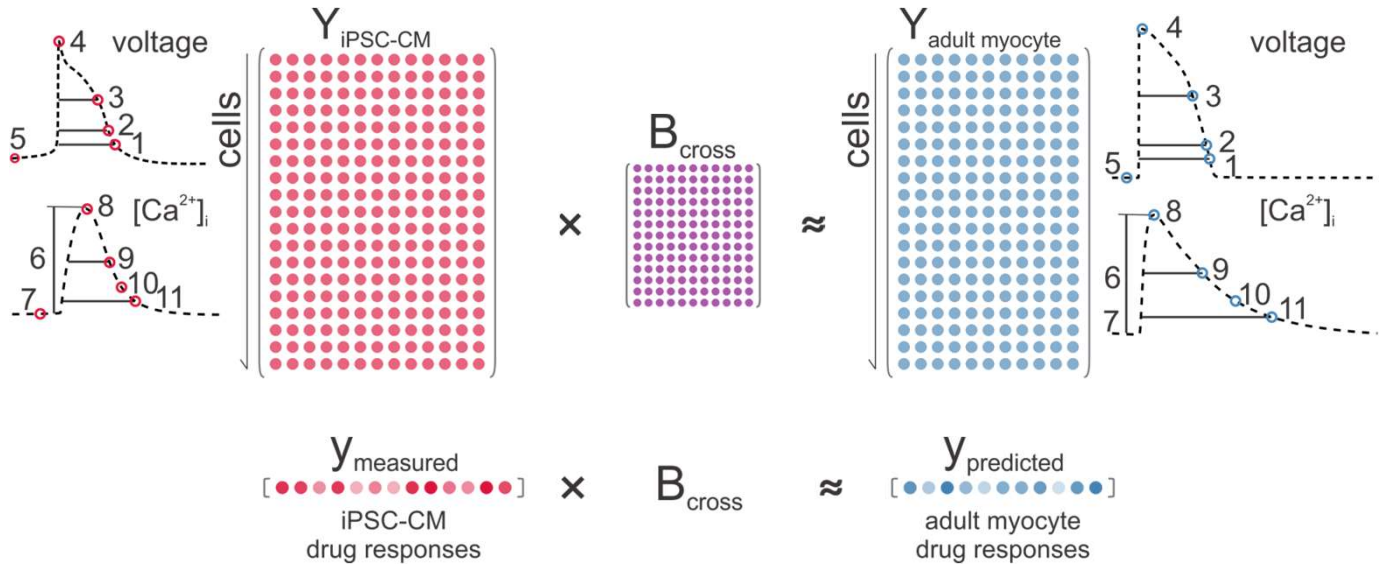
Population simulations and physiological metrics quantification were performed for both adult myocyte and iPSC-CM models

See, for instance:

Sobie (2009) *Biophys. J.* 96:1264-74.

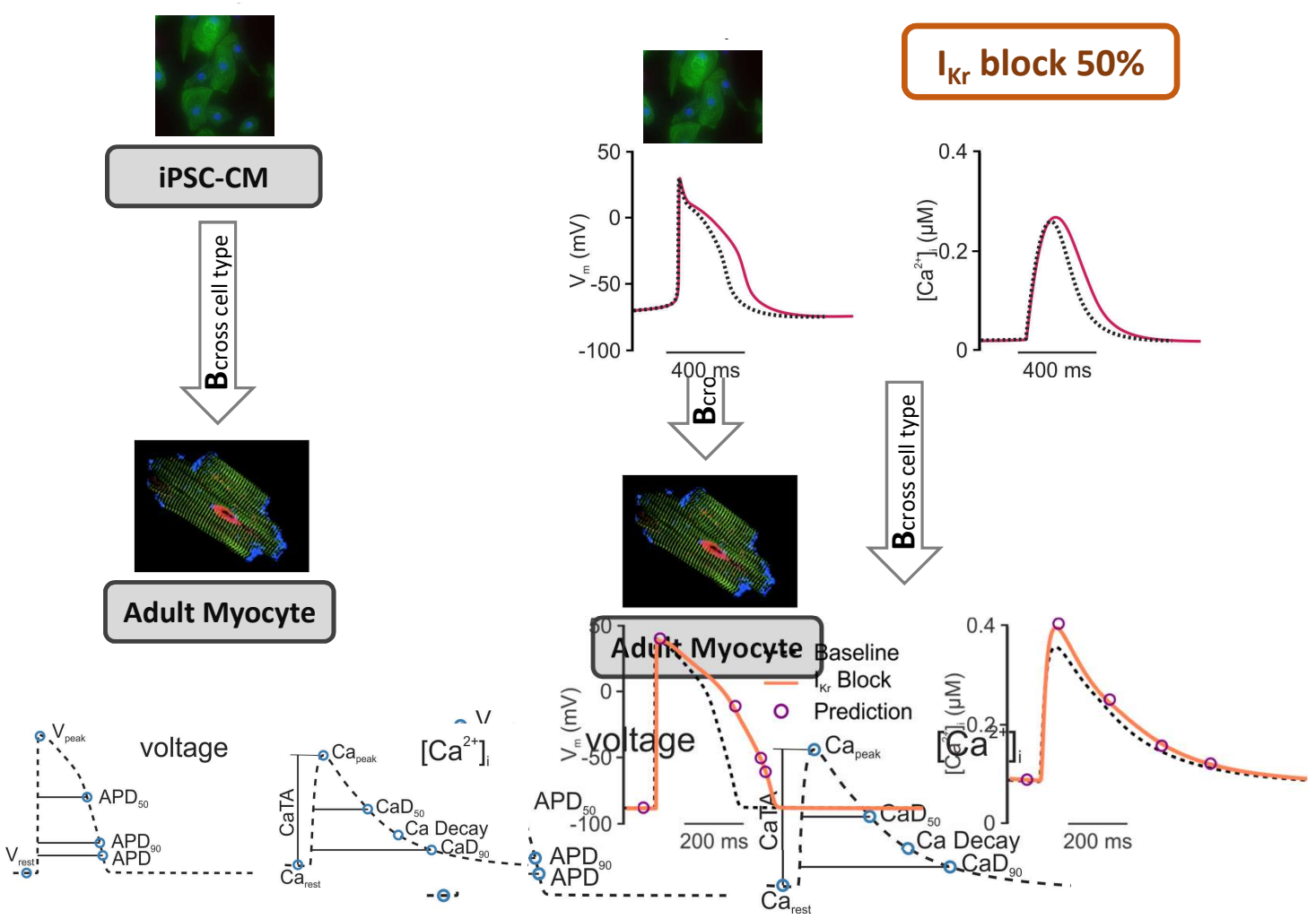
Sarkar & Sobie (2011) *Heart Rhythm* 8:1749-55.

Cross-cell type models to predict drug responses



Can this approach quantitatively translate drug responses across cell types to overcome experimental model limitations?

Does the regression model predict drug effects?

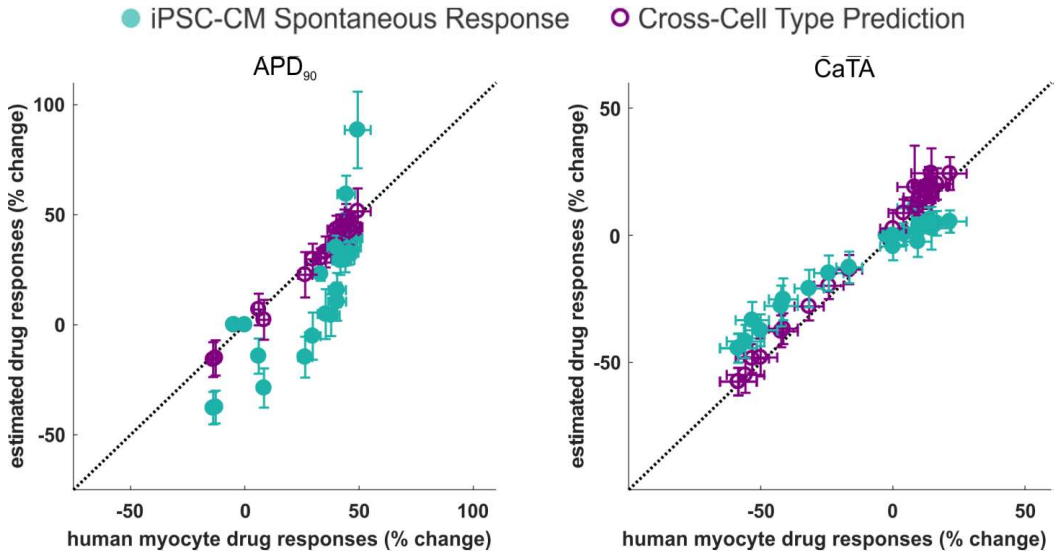


Gong & Sobie *npj Systems Biology & Applications* (2018) 4:11.

Regression model corrects for mismatches

Clinical drugs were simulated in a 100 cell heterogeneous population
30 drugs, block of 6 ion channels as recently measured

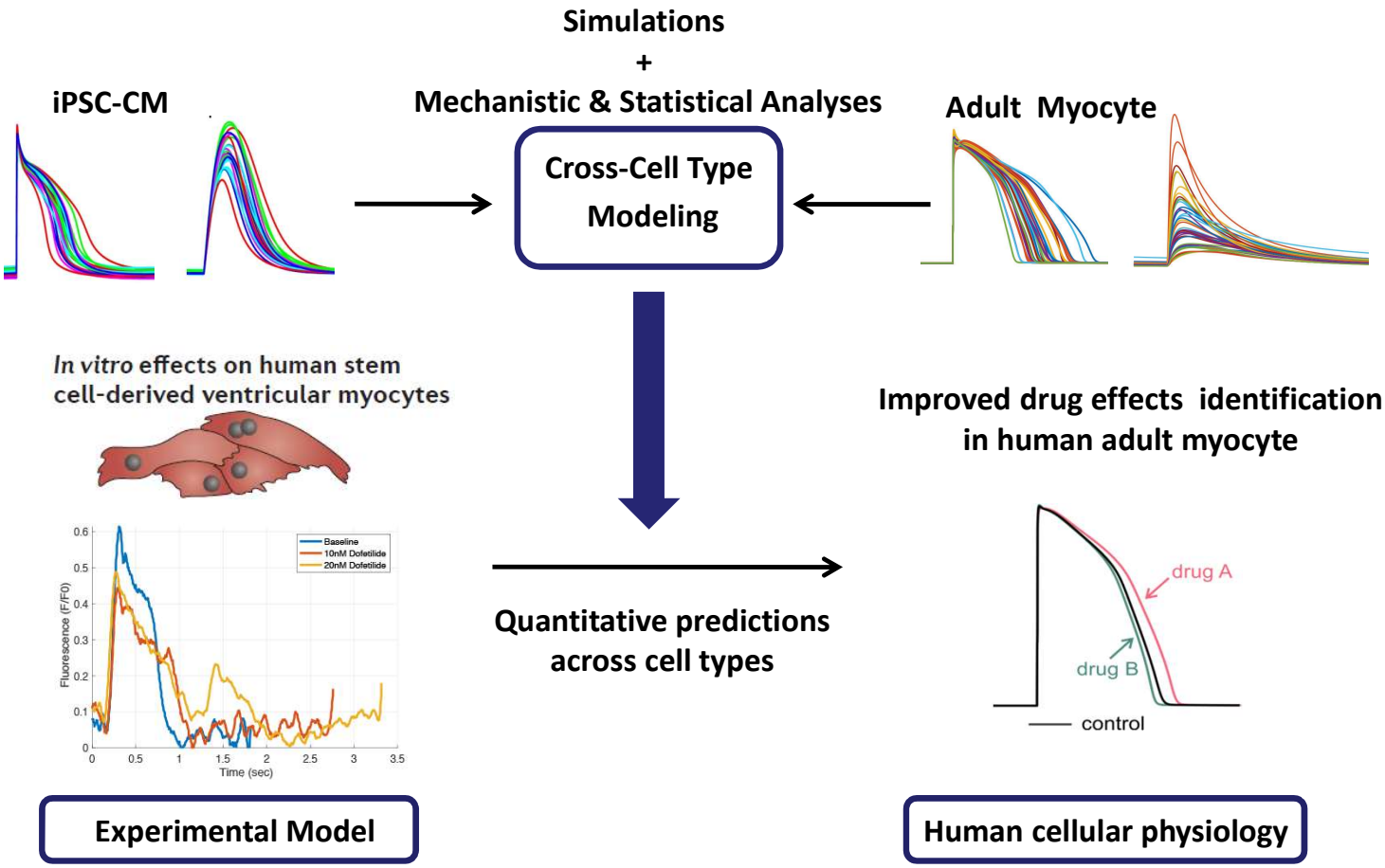
Crumb et al. *J Pharm. Tox. Meth.* (2016) 81:251-262.



Cross-cell type modeling can identify and correct mismatch of complex drug effects between cell types

Gong & Sobie *npj Systems Biology & Applications* (2018) 4:11.

Proposal: an enhanced role for modeling in CiPA



Outline

Computational prediction of proarrhythmia

Megan Cummins Lancaster, MD/PhD
Vanderbilt University

Cummins Lancaster & Sobie, *Clinical Pharmacology & Therapeutics* (2016) 100:371-379.

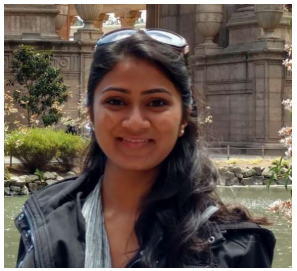
Predicting drug effects across cell types

Jingqi Gong
PhD candidate

Gong & Sobie *npj Systems Biology & Applications* (2018) 4:11.

Identification of susceptible sub-populations

Varshneya, Mei, & Sobie, unpublished work in progress



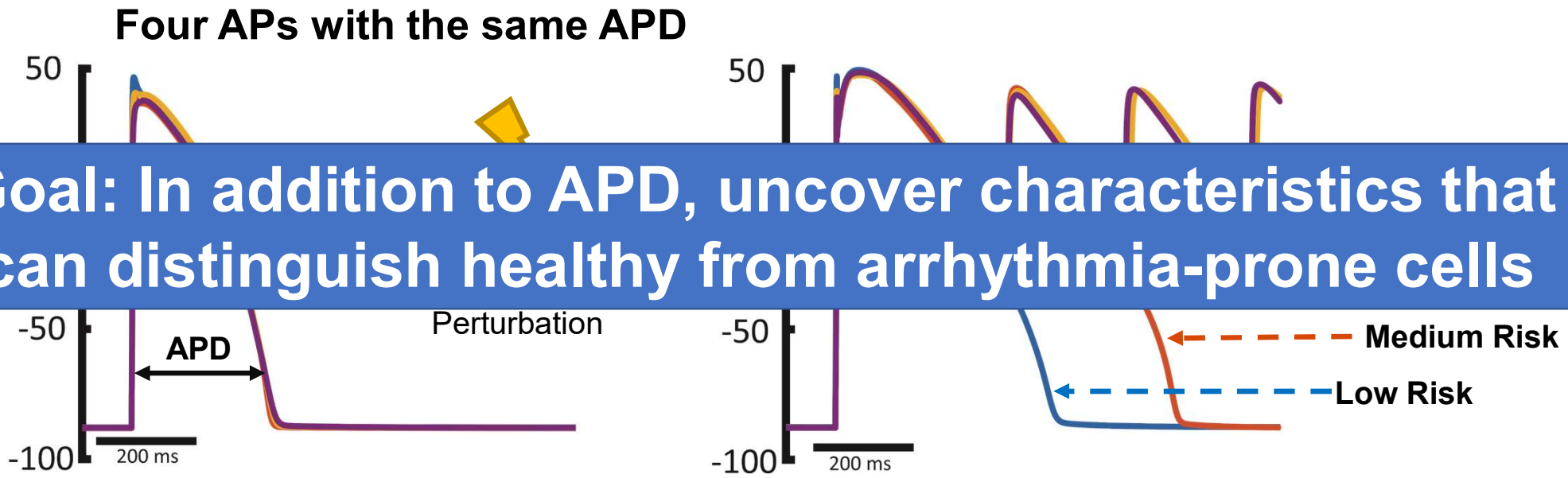
Meera Varshneya
PhD candidate

From a population of cardiomyocytes (patients), what experiments would you conduct to predict which individuals are susceptible to arrhythmias?

Most commonly, APD is measured and used to predict risk…

APs with longer APDs → more susceptible
APs with shorter APDs → less susceptible

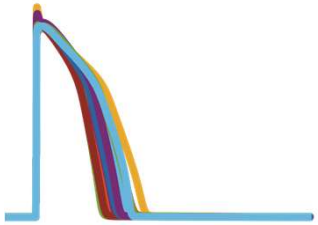
Cells with the same APD but different levels of susceptibility



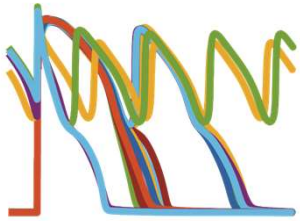
Goal: In addition to APD, uncover characteristics that can distinguish healthy from arrhythmia-prone cells

Research Strategy

1. Build large population of APs

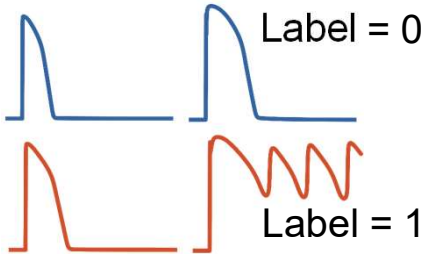


2. Apply Trigger

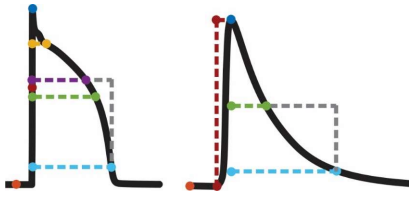


Increase I_{CaL}, I_{Kr} Block, Hypokalemia

3. Label APs based on EAD formation



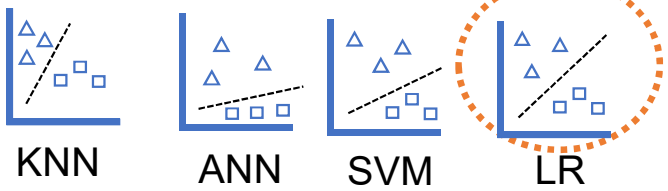
4. Measure parts of the baseline population



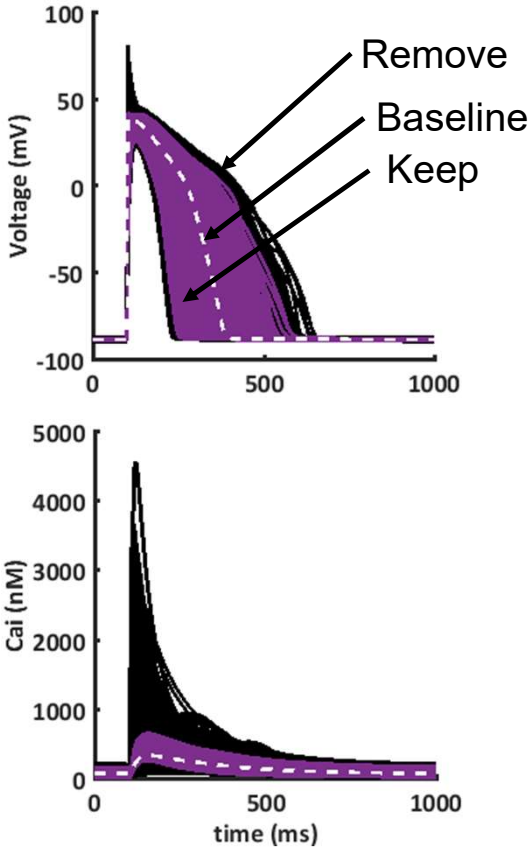
5. Apply various machine learning algorithms

KNN, SVM, Random Forest, Logistic Regression, Neural Network

6. Determine which algorithm made the best prediction



Calibrate population based on experimental data

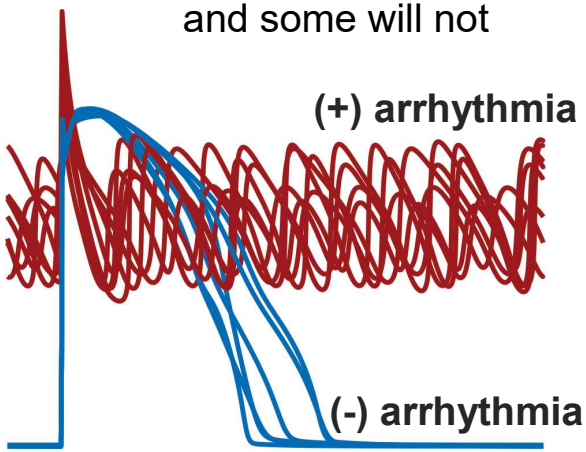


Passini et al. *J Mol Cell Cardiol* 2016

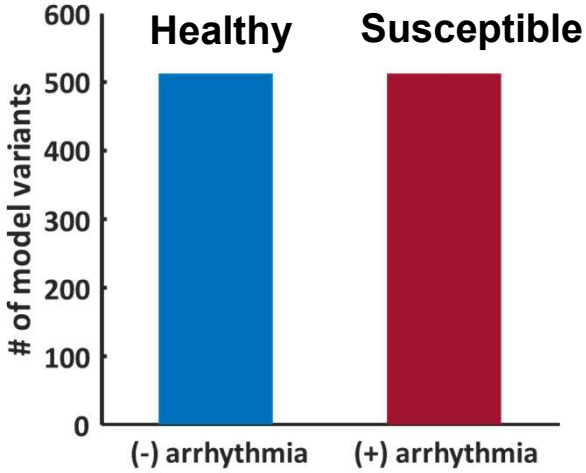
Apply Trigger on Population

Increase L-type calcium current

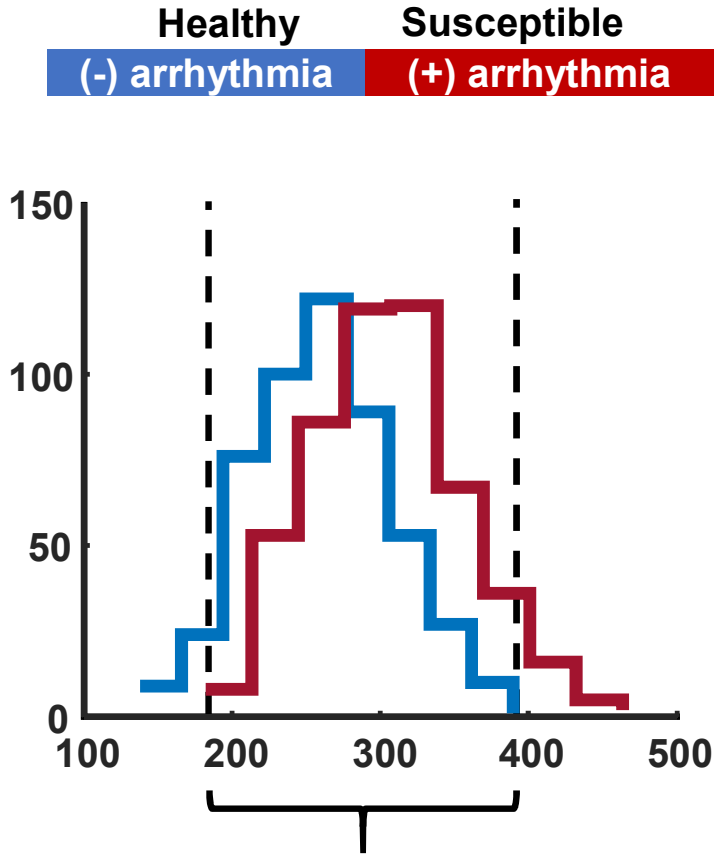
Some will form an arrhythmia, and some will not



Split Initial Population into Two Groups

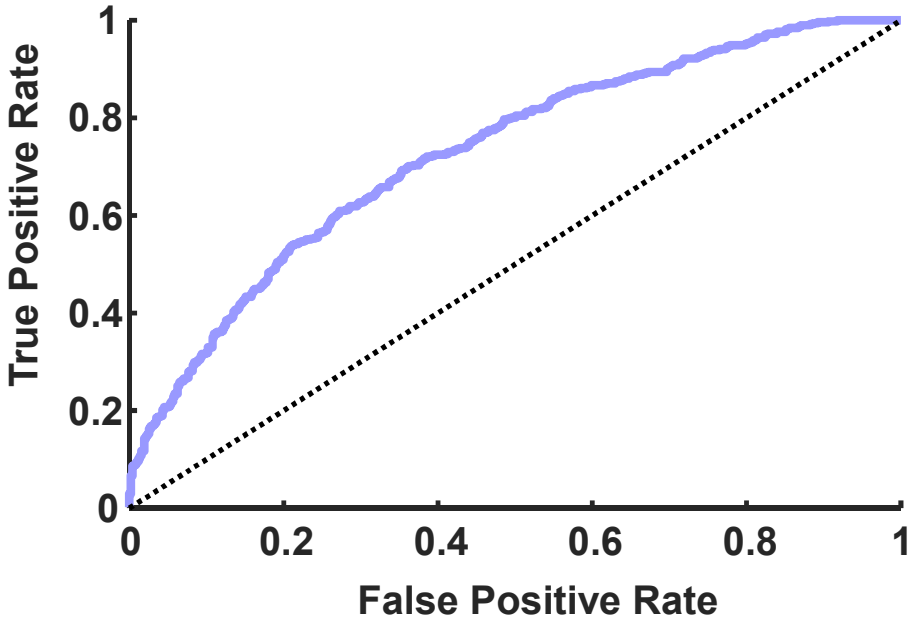


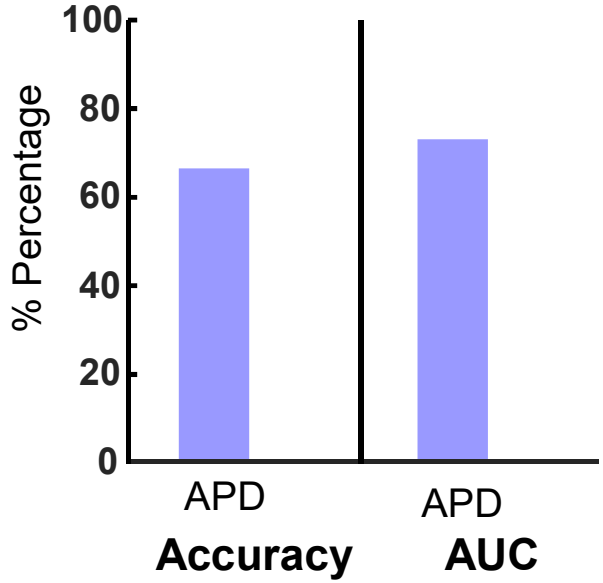
APD_{90} can separate groups, but there is room for improvement



Range of APD_{90}s where prediction could go either way

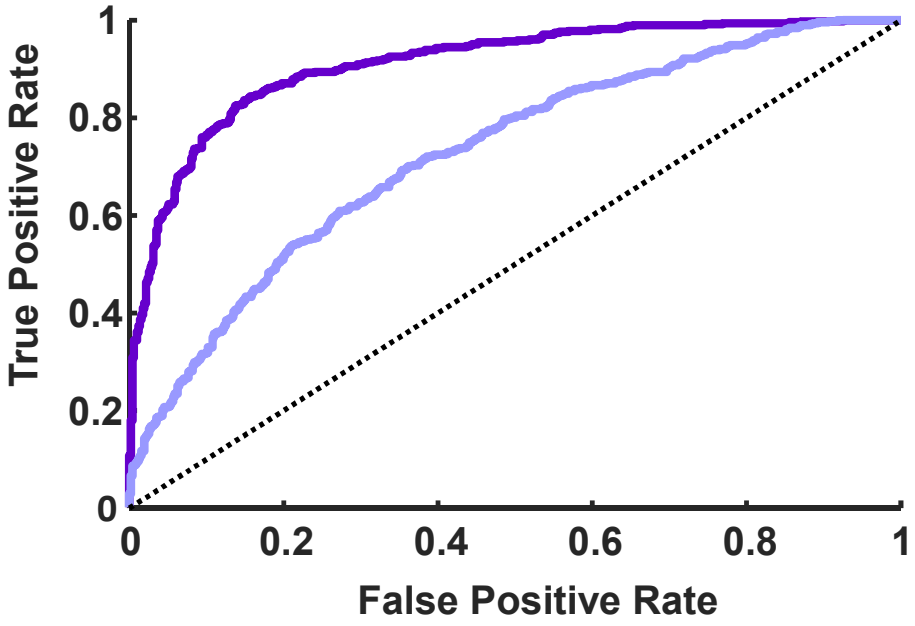
Machine learning with two features (APD_{50} and APD_{90}) is decent

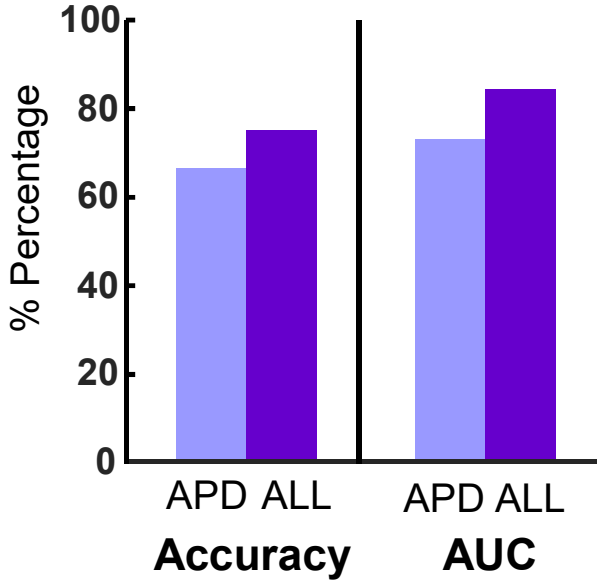




Results obtained using Support Vector Machine.

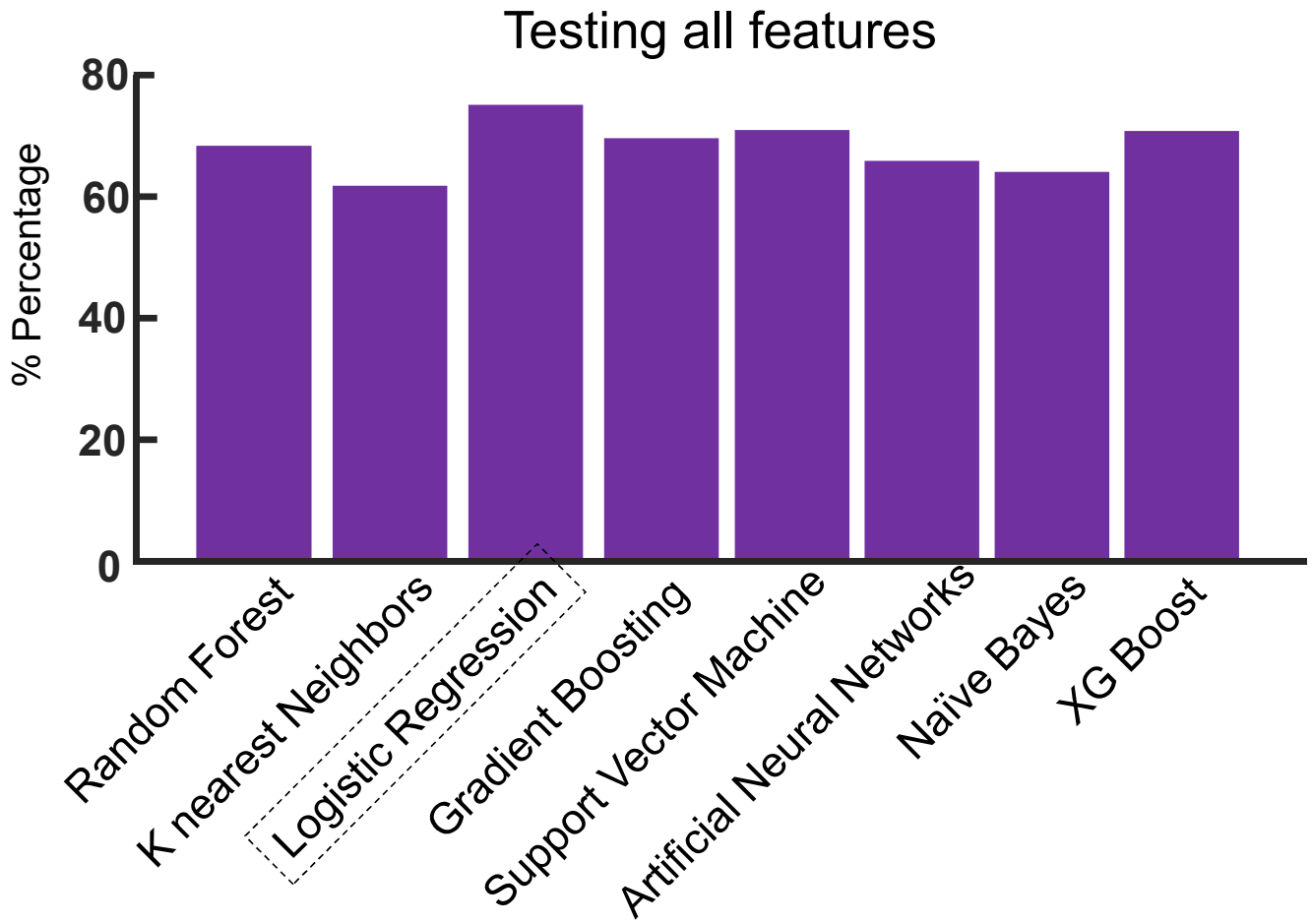
Machine Learning performance improves when additional features of AP and CaT are added





Results obtained using Support Vector Machine (APD) and Logistic Regression (ALL).

Accuracy is similar between various machine learning algorithms

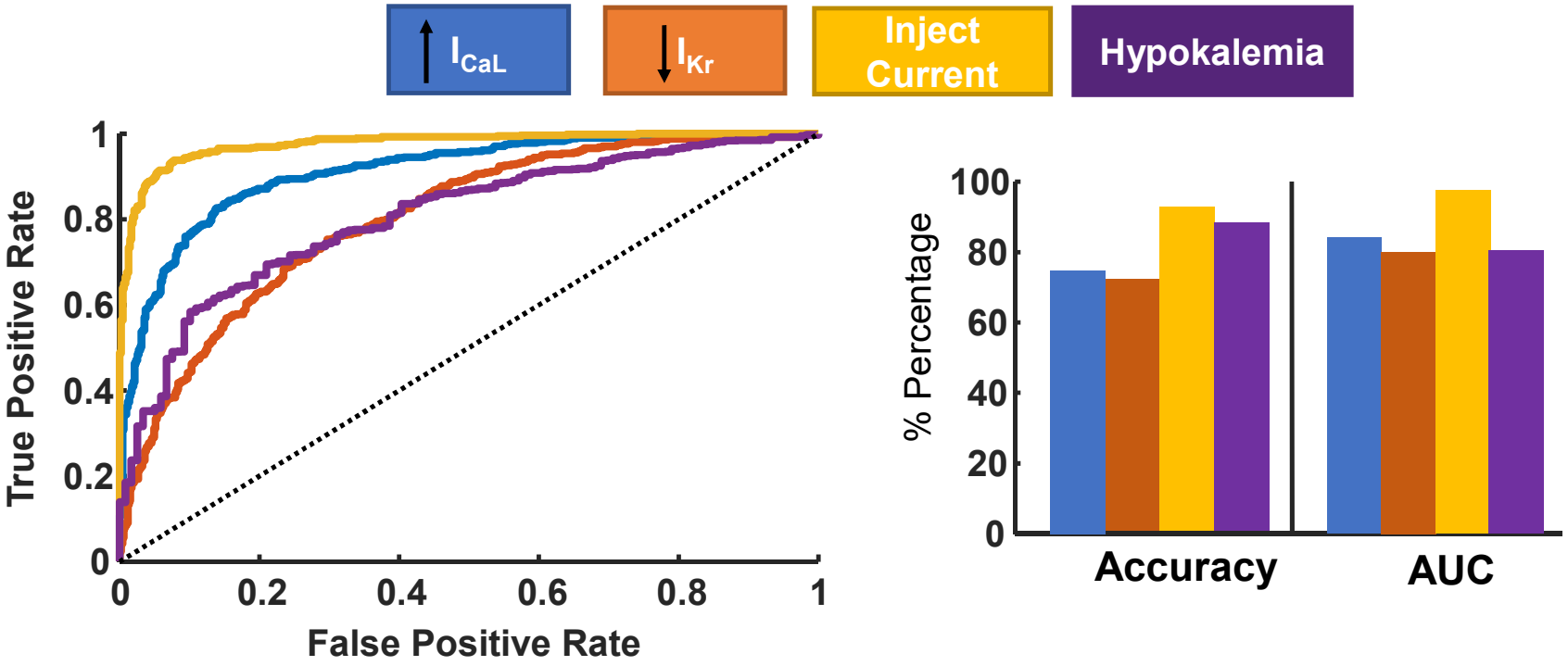


Questions to be addressed…

- Does machine learning performance remain the same no matter what trigger is applied?
- Does machine learning performance change based on the parameters that are varied in the model?
- What other features about the baseline population can be calculated to improve machine learning performance?

Does machine learning performance remain the same no matter what trigger is applied?

Machine learning performance changes based on the applied trigger.

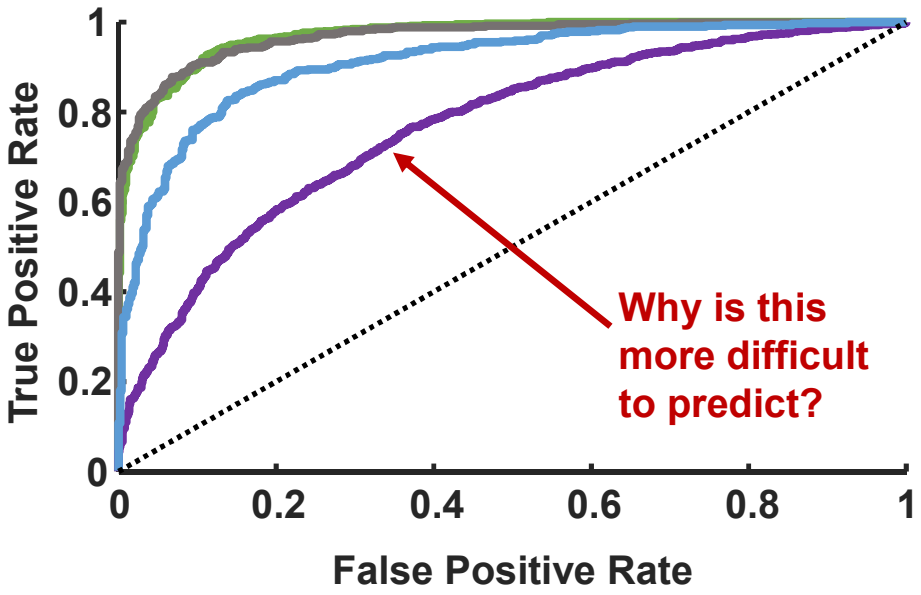


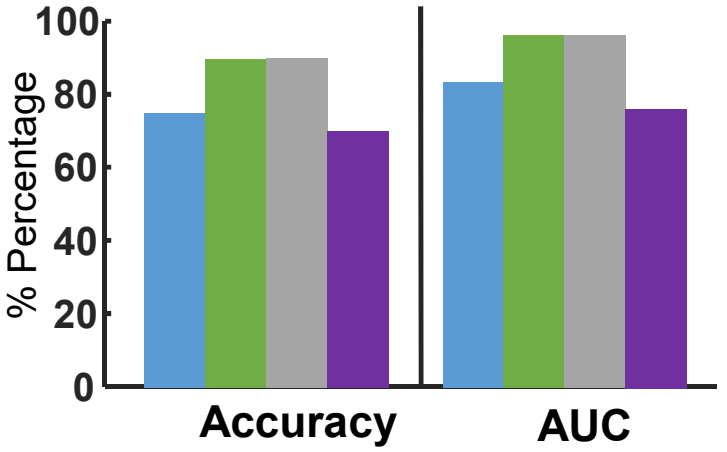
Results obtained using Logistic Regression. All features were used.

Does machine learning performance change based on the parameters that are varied in the model?

Machine learning performance changes based on the parameters that are varied.

Vary maximal density | Vary kinetics | Vary voltage dependences | Vary everything



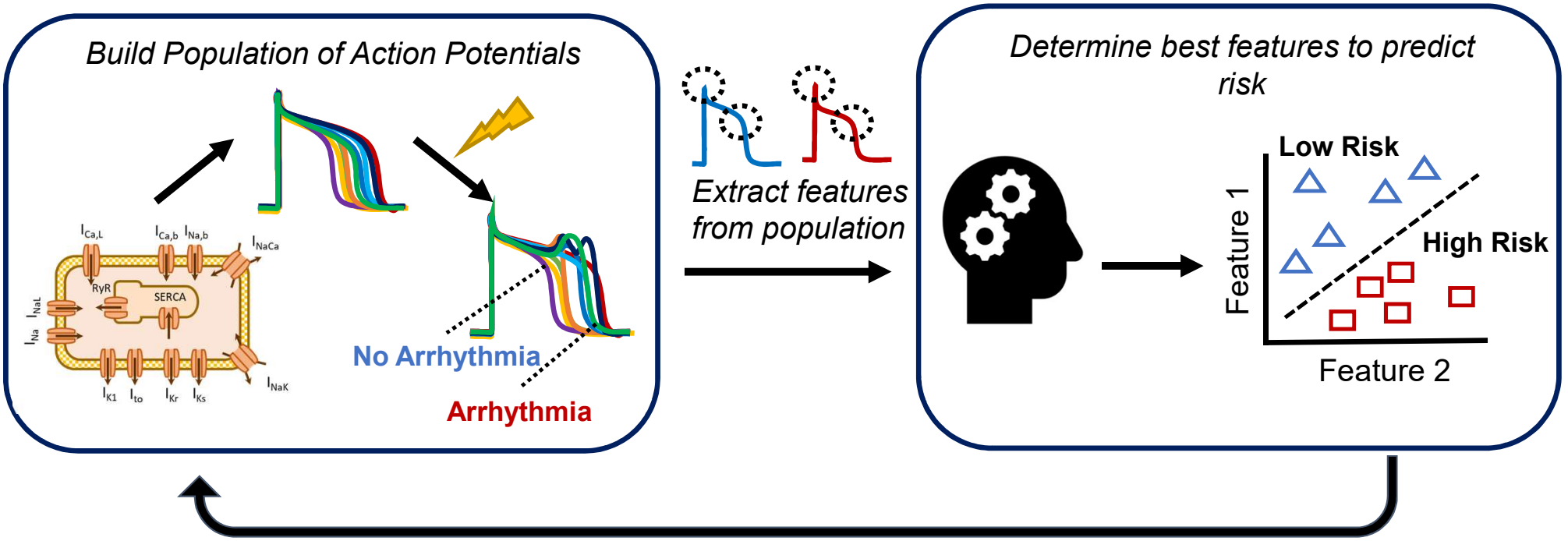


Results obtained using Logistic Regression.

What other features about the baseline population can be calculated to improve machine learning performance?

QUANTITATIVE SYSTEMS PHARMACOLOGY

MACHINE LEARNING ALGORITHMS



Accuracy ~75%

Can we return to the systems pharmacology models to engineer better features and improve performance?

Acknowledgements

Current lab members:

Jingqi Gong
Jaehee Shim
Chiara Campana
Meera Varshneya
DeAnalisa Jones

Recent alumni:

Rafael Dariolli, PhD
Ryan Devenyi, MD, PhD
Tobias Holden
Jake Hartman-Kenzler
Megan Lancaster, MD, PhD

Collaborators:

David Christini, Weill Cornell
Beatriz Trenor, Valencia
Ravi Iyengar, Mount Sinai

National Heart Lung and Blood Institute
National Institute of General Medical Sciences
NIH LINCS PROGRAM
NSF
American Heart Association
Learn and Live