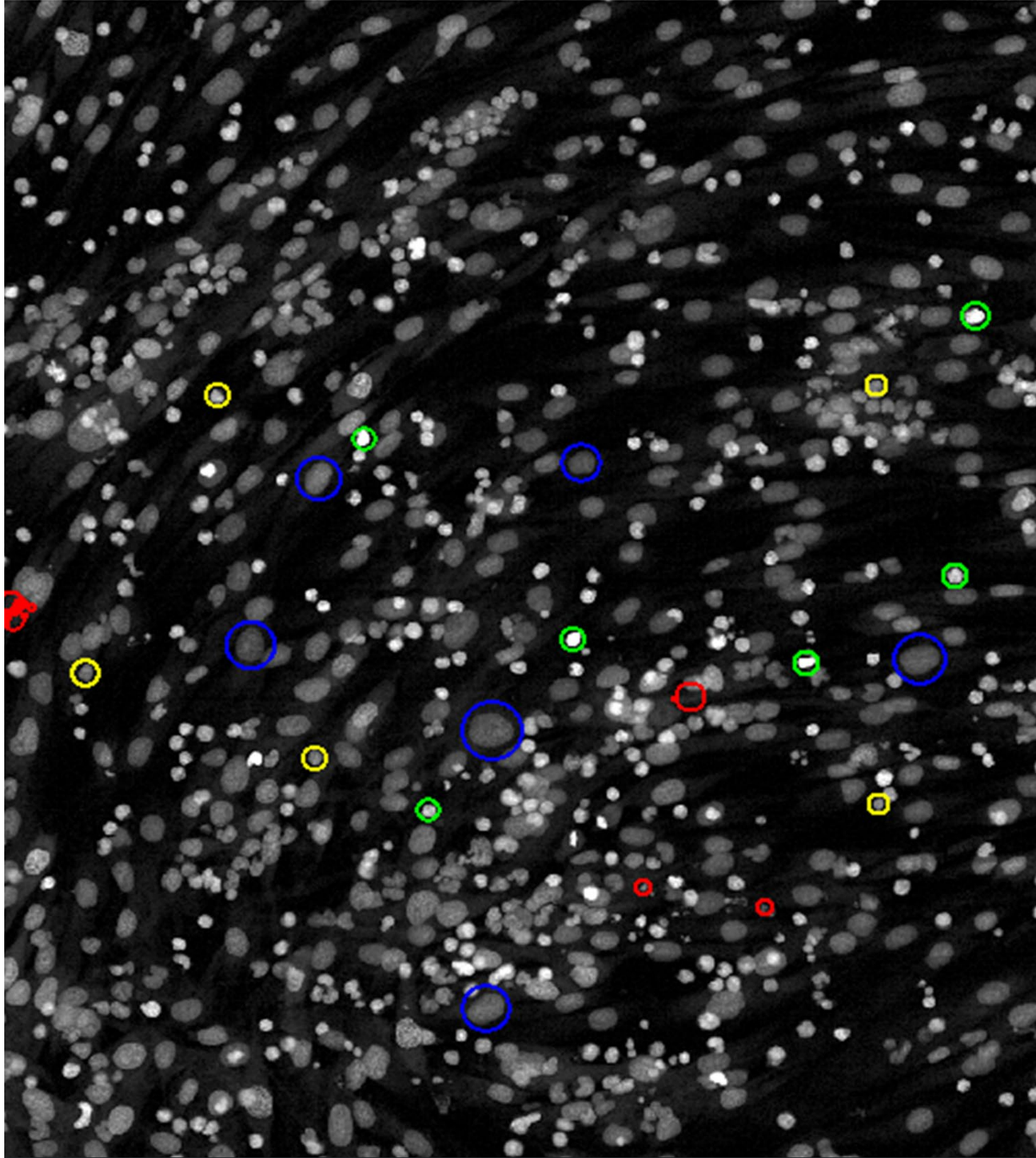
Integrating Robust Machine Learning Classifiers to Automate Fluorescence Imaging-based *Ex Vivo* Drug Sensitivity Testing of Acute Lymphoblastic Leukemia

Landon Choi, MS

Sr. Researcher, Department of Pharmacy and Pharmaceutical Sciences

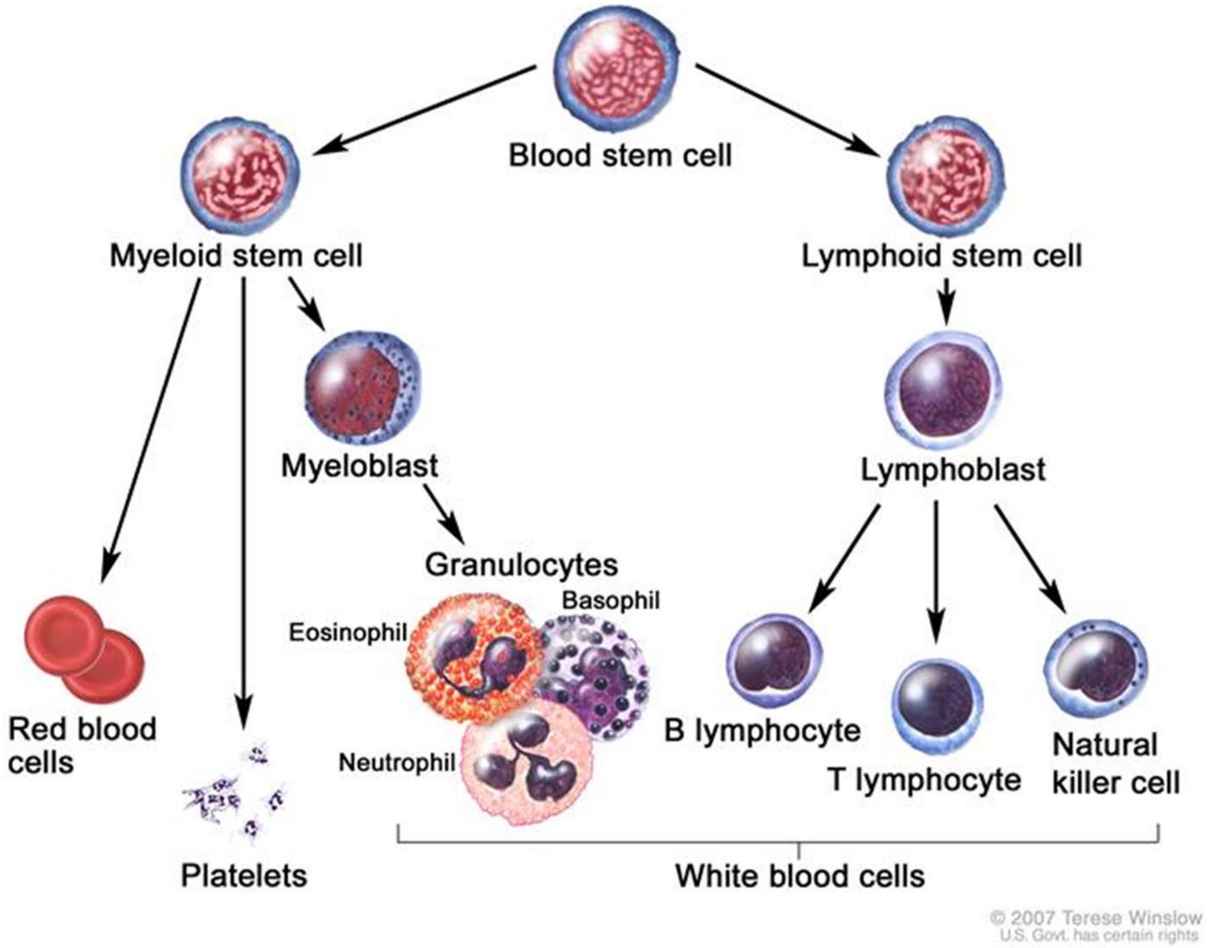
09/27/2024



Affiliations & Disclosures

- St. Jude Children's Research Hospital
- No Disclosures

Background Information



Population

Pediatric Acute Lymphoblastic Leukemia (ALL) is the most common type of childhood cancer representing ~25% of all new pediatric cancer diagnoses within the US.

Leukemia

Two main Types; B-ALL and T-ALL depending on which lymphocyte the disease originates from.

- 23 Subtypes for B-ALL
- 12 Subtypes for T-ALL

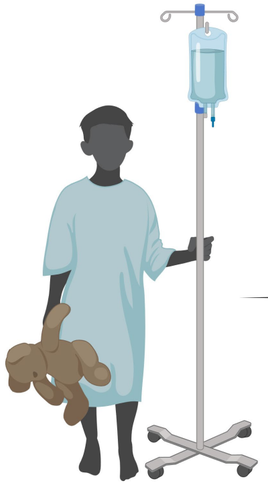
Pharmacotyping

Pharmacogenomics is the study of how genetic attributes affect drug response; pharmacotyping is defining a patient's leukemia blast phenotype *in vitro*.

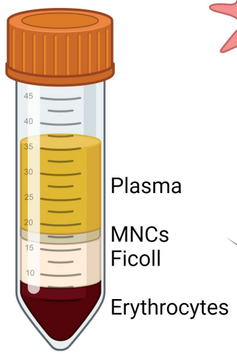
Our pharmacotyping assay is reliant on fluorescent imaging and a Mesenchymal Stromal Cell (MSC) co-culture to evaluate a sample's morphology & sensitivity.

Brady S, et al. *Nat Genetics* 2022 & SEER 22 Data (2017-21)

Pharmacotyping



BM or PB sample received from patient with ALL



Ficoll density gradient centrifugation to isolate leukemia cells

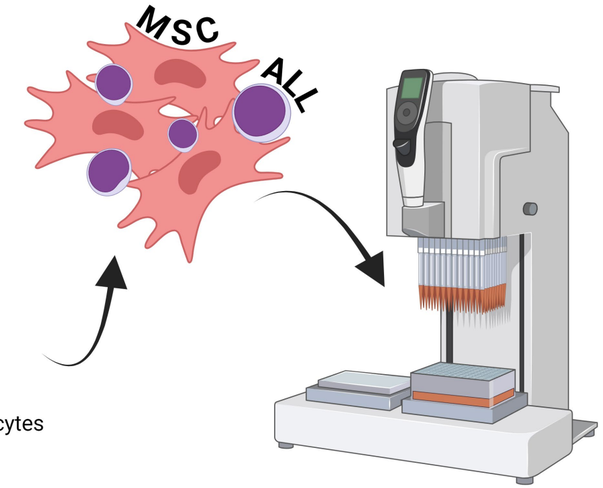
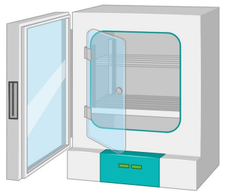
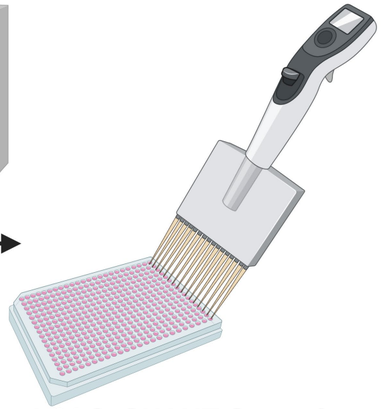


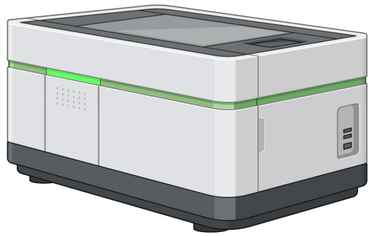
Plate leukemia cells and drug dilutions on 384 well plate with MSCs



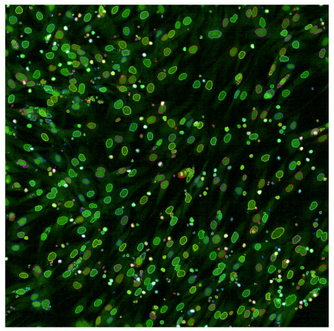
96 hours



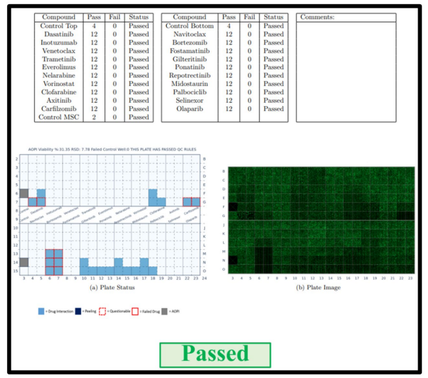
Add CyQUANT Green & AOPI followed by 15 minute incubation



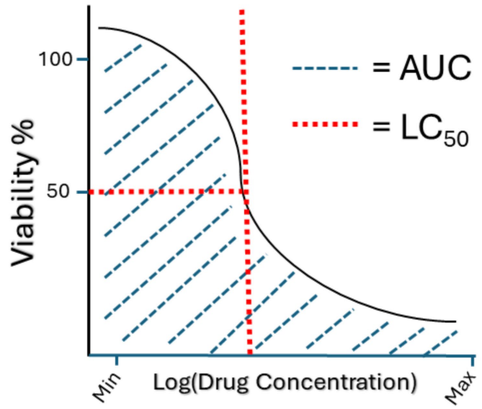
in vitro fluorescent imaging



Imaging pipeline analysis



Quality Control pipeline analysis

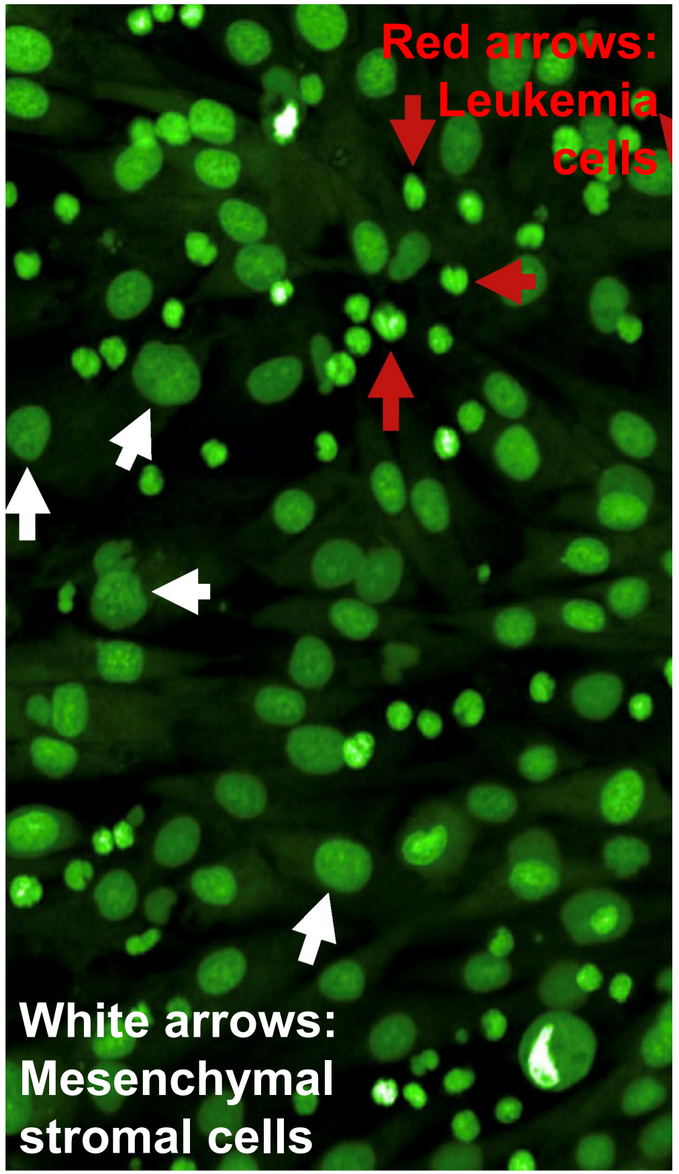


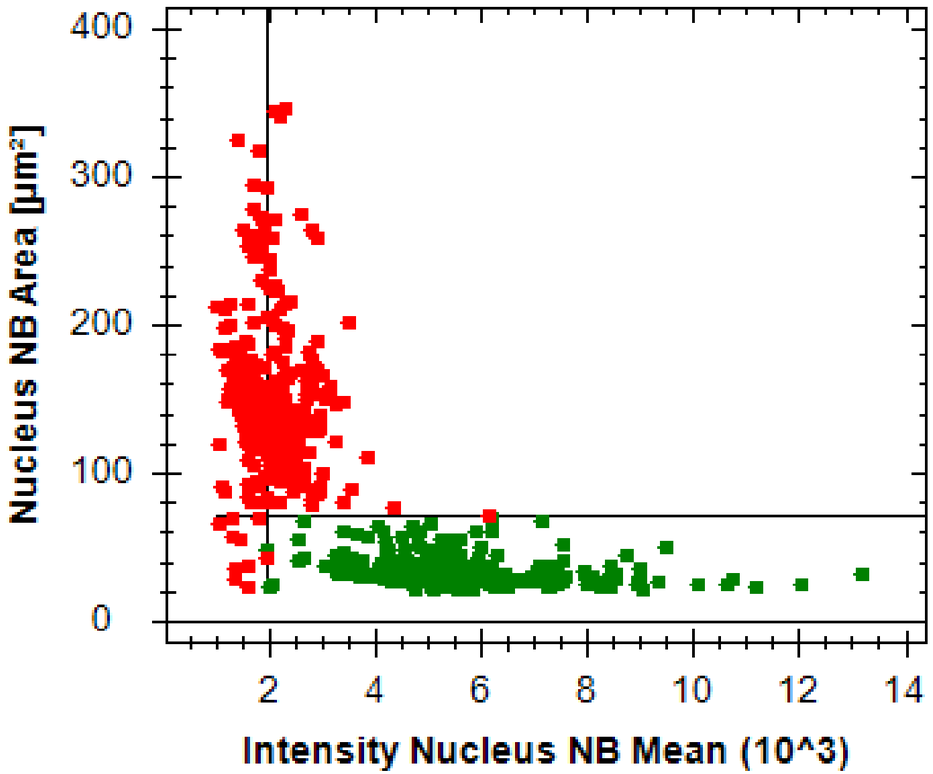
(AUC): Area Under the Curve

(LC50): Lethal Concentration when viability reaches 50%

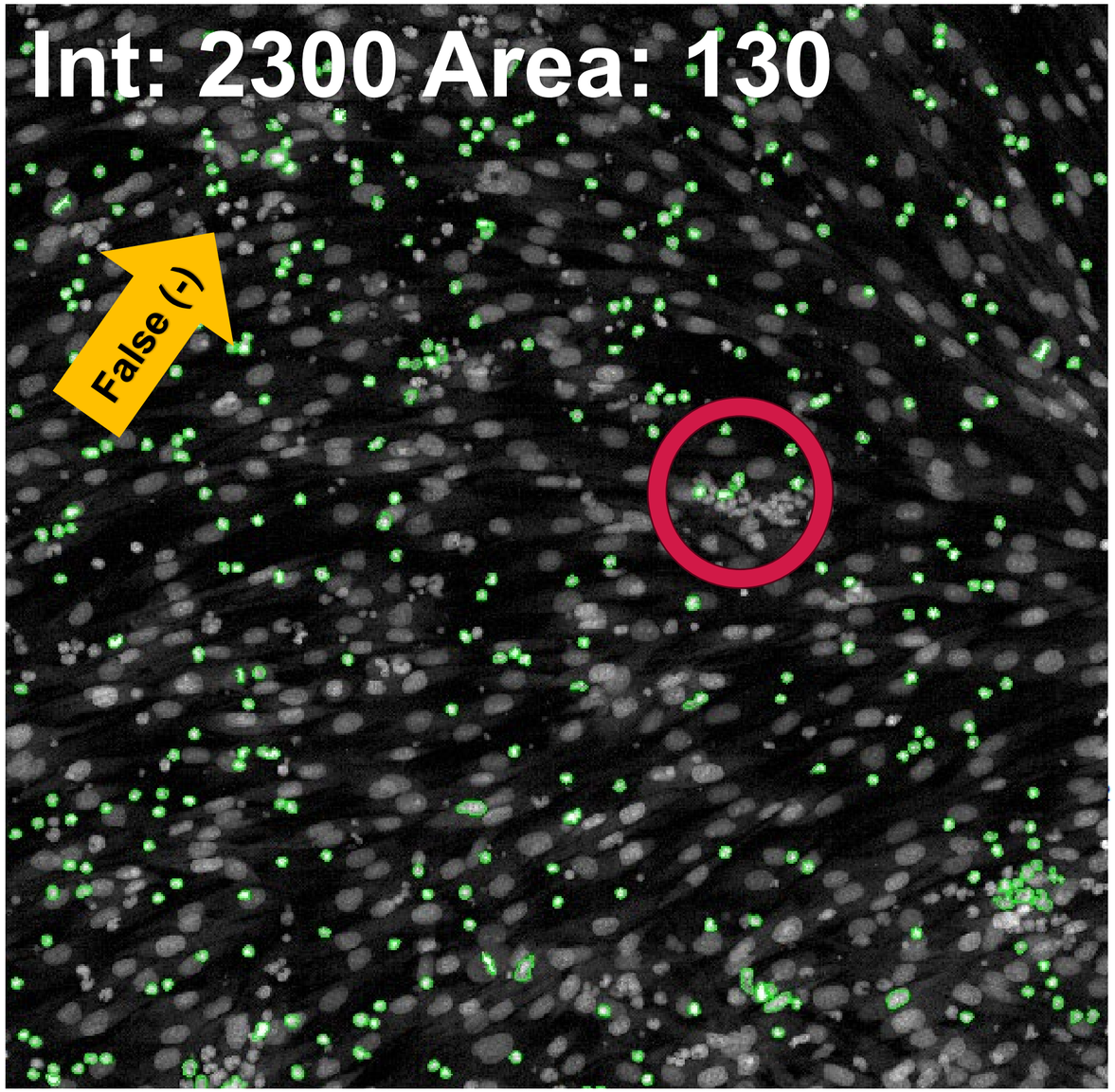
Created with BioRender.com

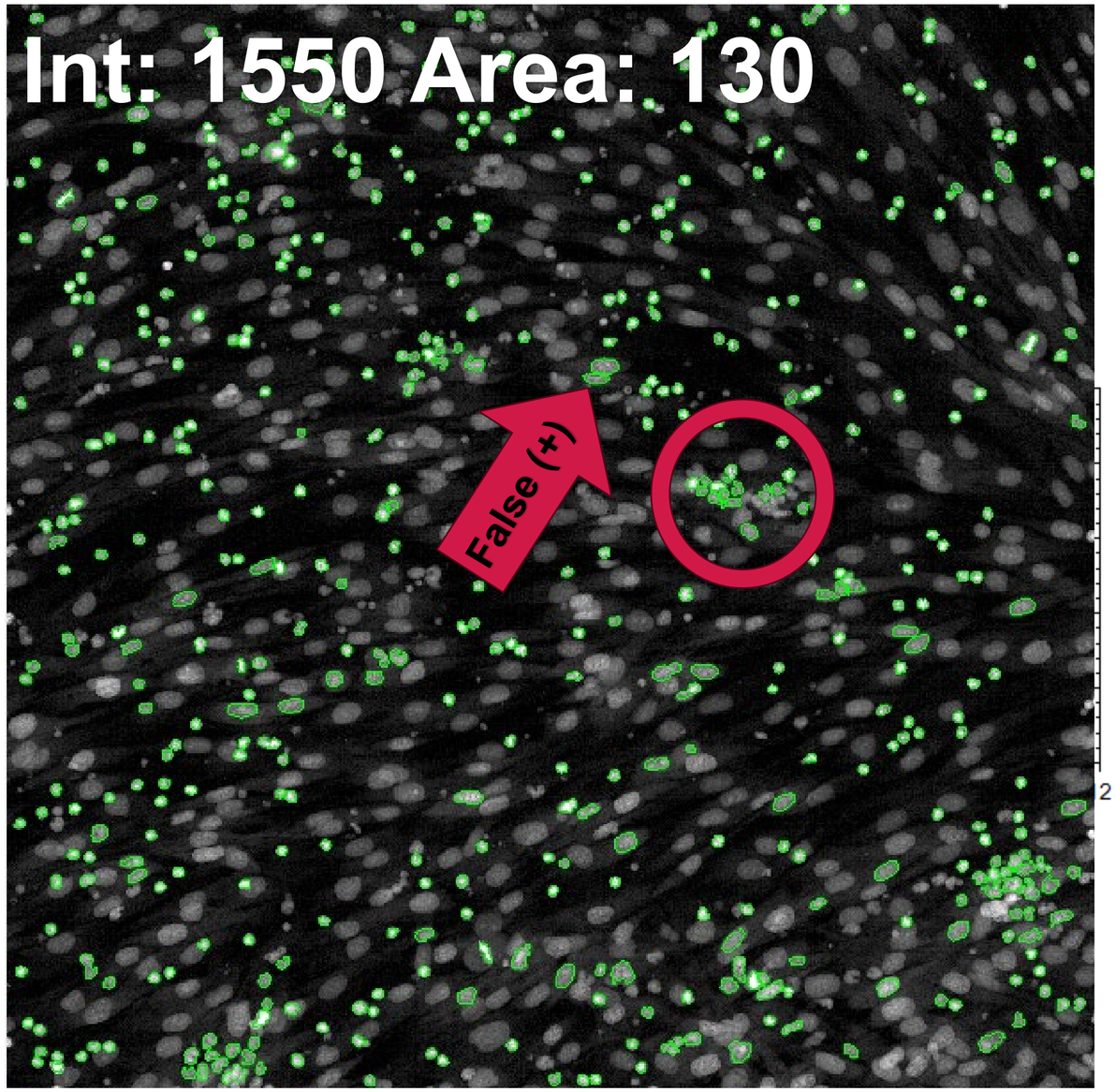
Objective: Discerning leukemia cells from MSCs using nucleus area and intensity.





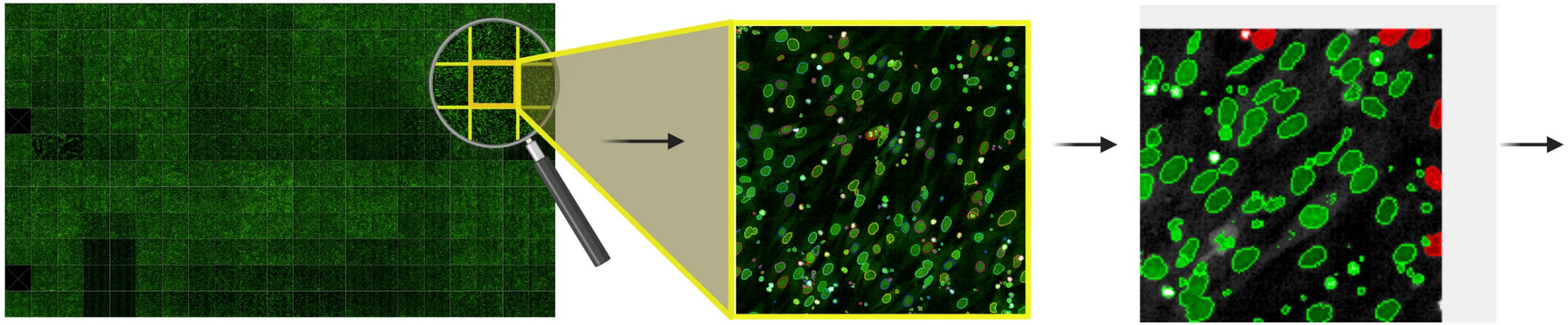
Manual Thresholding Pipeline





Same Well, Field of View, and Sample

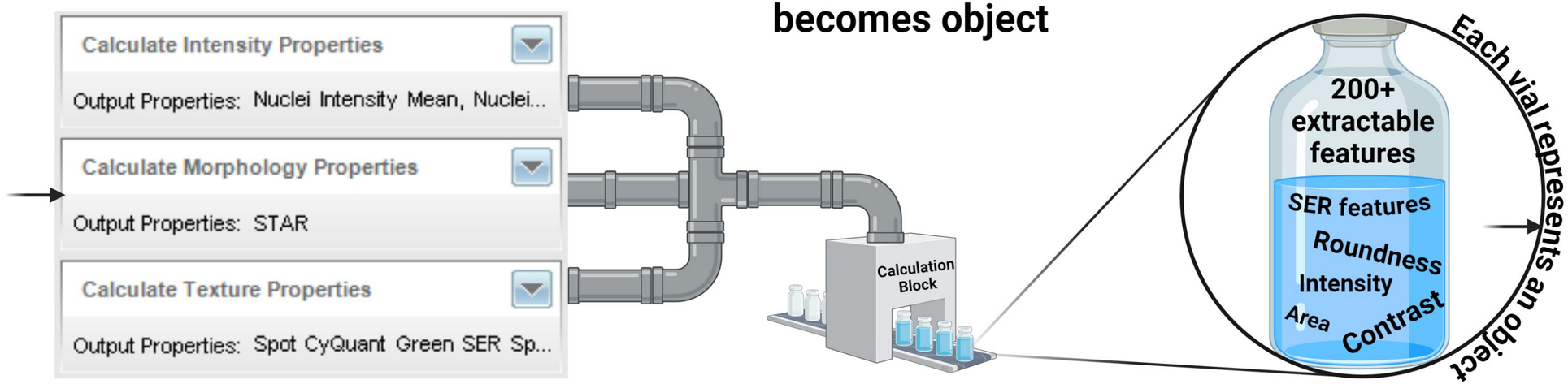
Machine Learning Pipeline Workflow

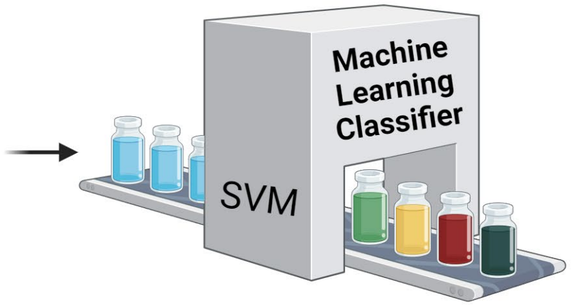


Input Image:
1 well = 9 FOVs

Object Identification:
Fluorescent cell/artifact becomes object

Refine Population:
Remove Border





Linear Classifier uses training to label objects into different populations.

Created with BioRender.com

Machine Learning Training

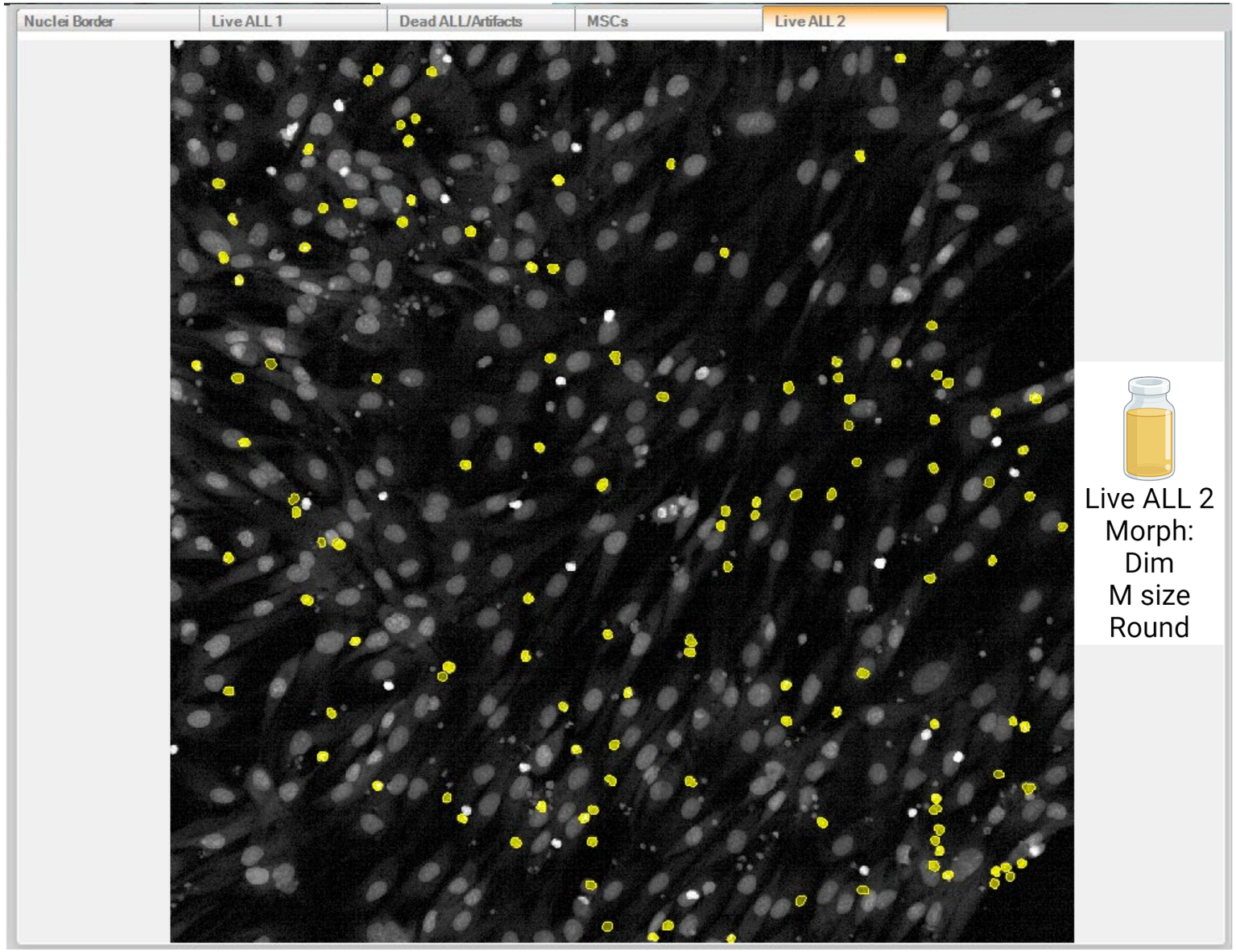
DATA SET INFO:

- Current model trained on 10 cases (DS0) (All cases passed QC)
- Objects selected from control wells [random], high concentration wells [random], low concentration wells [random], middle concentration wells [random].

Priority cell populations for model (n=100):

- Dead/Artifact
- MSCs
- Live ALL 1
- Live ALL 2

Viable ALL = Live ALL1 + Live ALL2
(Brighter ALL) (Dimmer ALL)



Differences in manual vs machine learning pipelines

Manual Thresholding

- Due to heterogeneity of samples, human input is required to determine the threshold based on intensity & area.
  - Potentially introduces bias.
  - Less consistent.
  - Classifies only 2 populations (MSC vs leukemia cell) therefore cannot label a cell as dead, etc.
  - Faster analysis time, due to less data.

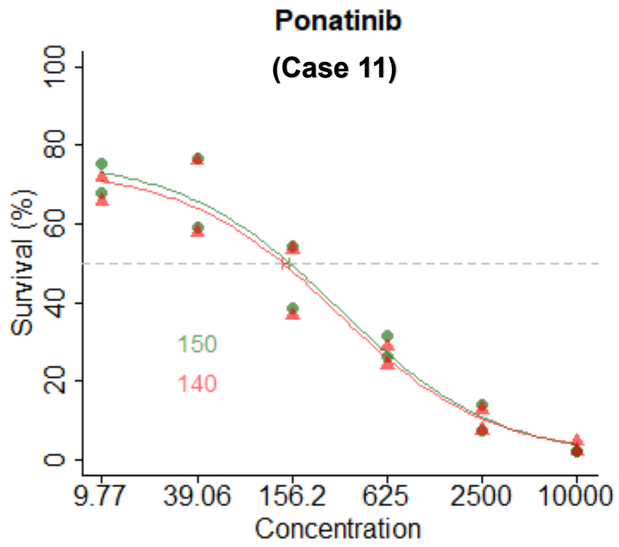
Pre-trained Machine Learning

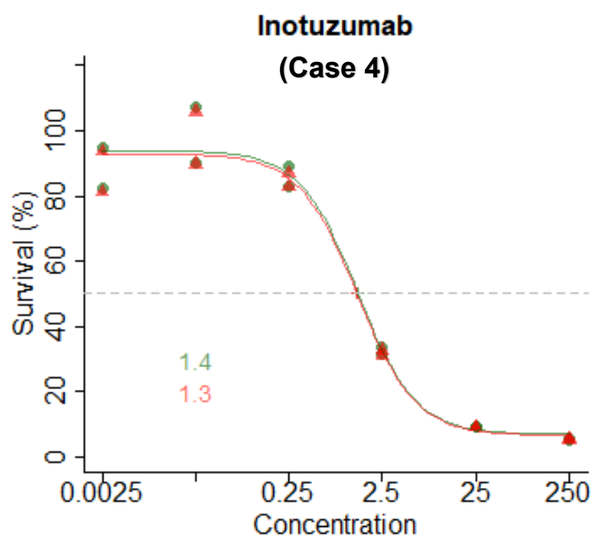
- Requires no human input for thresholding.
  - Binning of desired population to account for the variability among cases.
  - Uses Find Spots as its object identification block (as opposed to Find Nuclei/Cells).
  - Uses 200+ features to classify an object.
  - More populations for classifications reduces false positives.
  - Uses 3 different types of data collection blocks to make the classification.
  - Longer analysis time, due to more data.

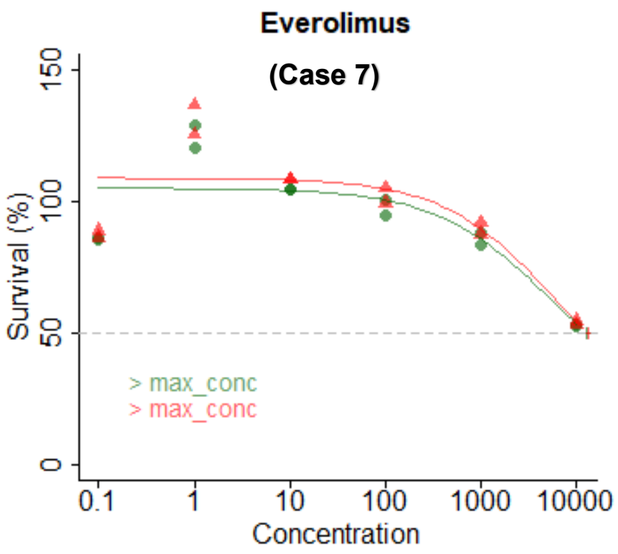
Validation of Machine Learning Pipeline

Comparison of LC50s

- Manual vs. Machine Learning
- Red = Machine Learning
- Green = Manual

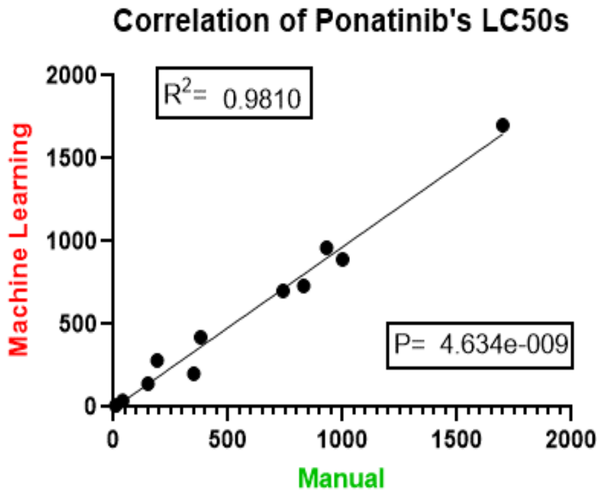


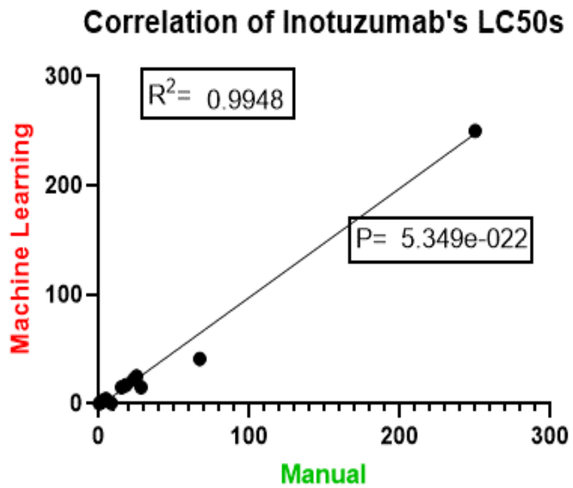


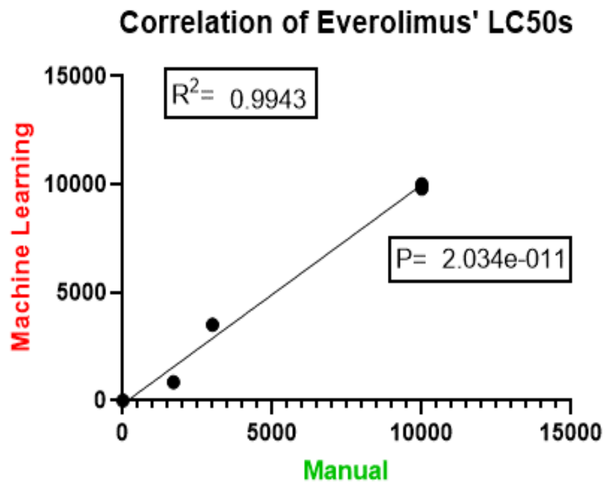


Correlation Plots

- Comparison of the imaging based LC50's between manual and automatic pipelines.
- Manual vs. Machine Learning



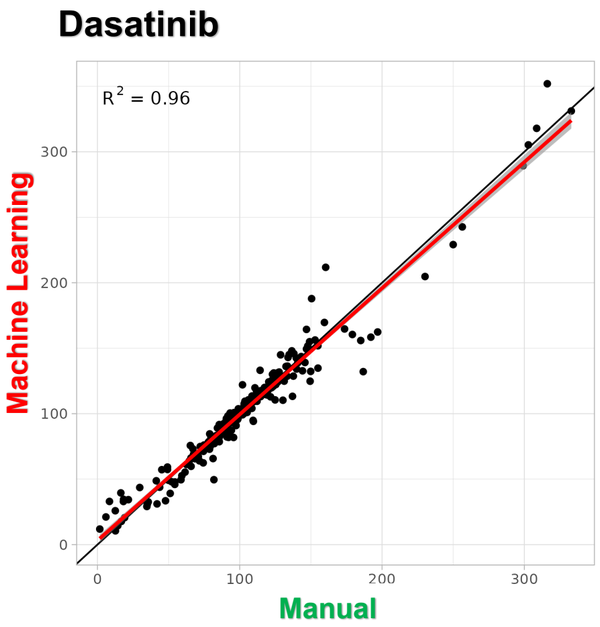


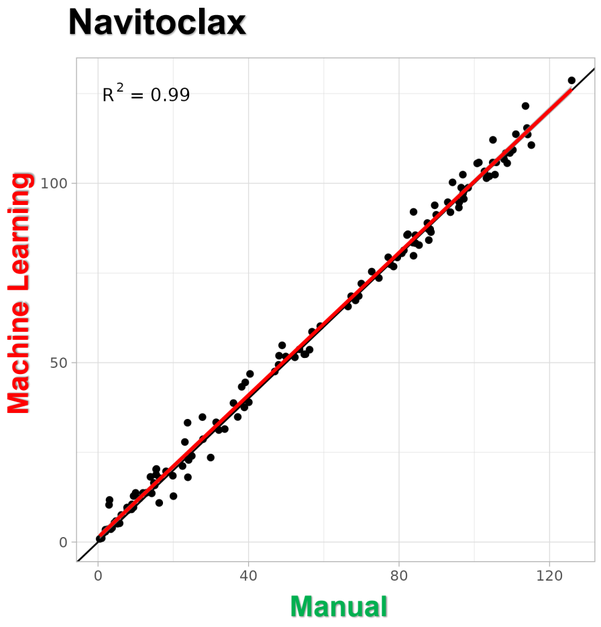


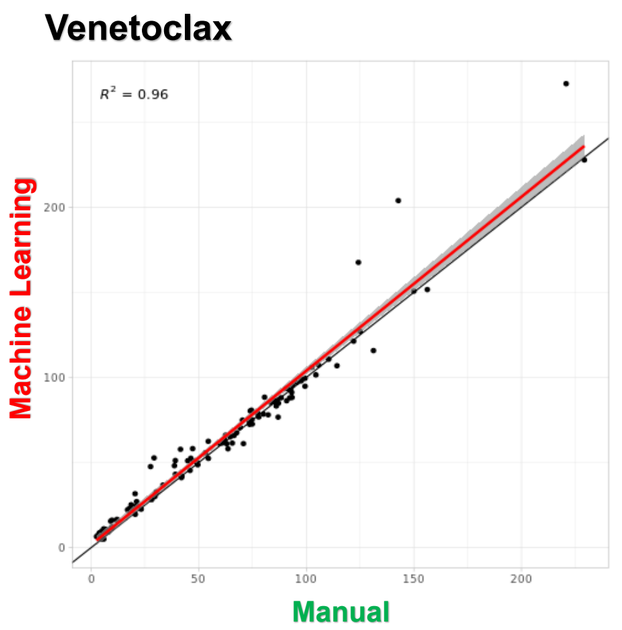
Validation of Machine Learning Pipeline

Comparison Plots

- R^2 = Correlation
- Black line R^2 = 100% correlation
- Manual vs. Machine Learning

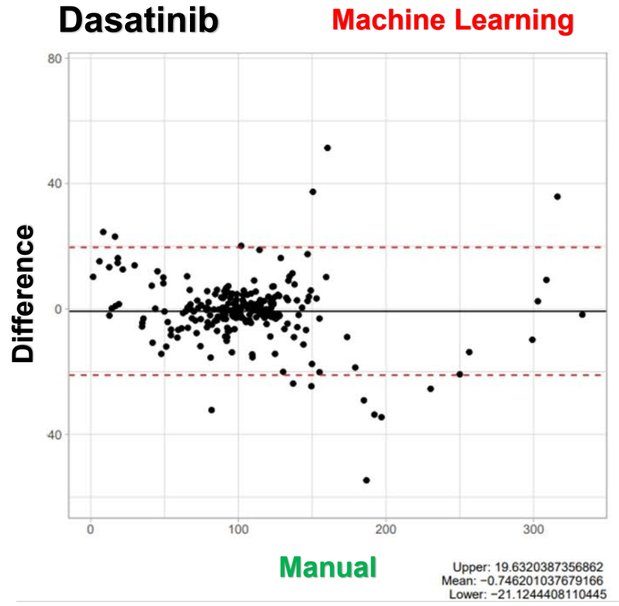


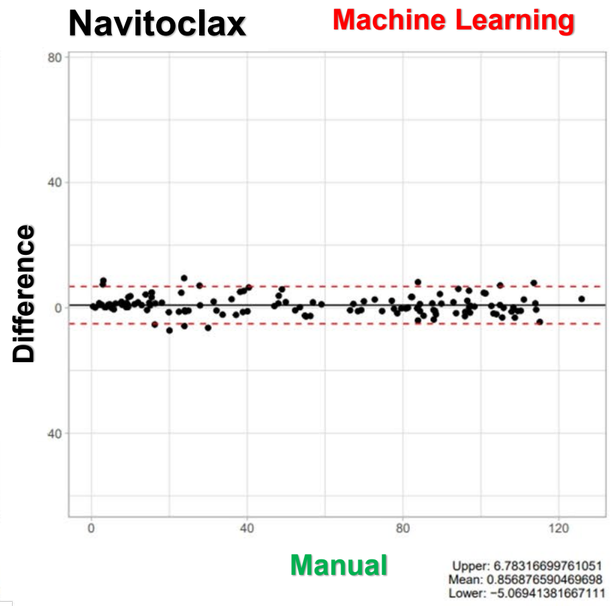


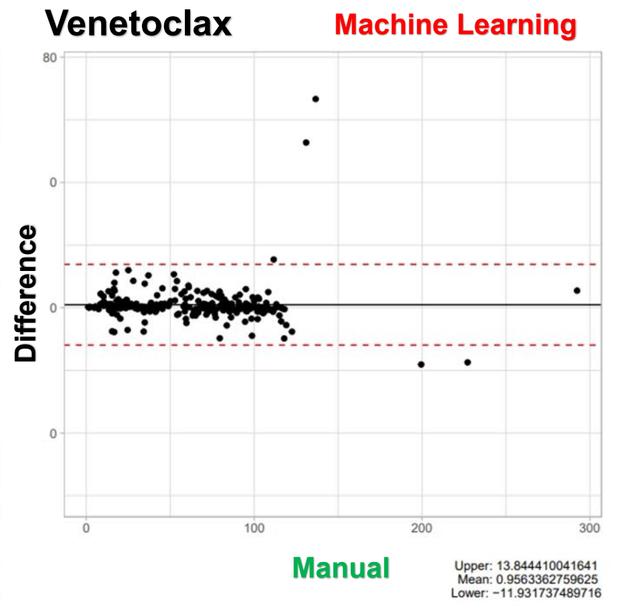


Bland-Altman Plots

- Final validation comparing the imaging based absolute percent viability differences.
- Manual (<0) vs. Machine Learning (>0)



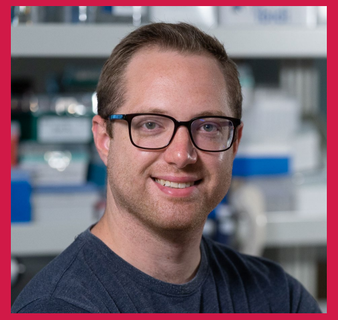




Key Take Aways

- Use of machine learning classifiers shows capability of discerning leukemia cells from MSCs using 200+ features.
- Machine learning classifiers provide comparable results to the manual method while enhancing consistency across heterogeneous samples.
- Use of machine learning classifiers decreases the chances of potential bias by removing human dependency (for discerning populations).

Pharmacotyping Team



Tony Brown, Ph.D.

Amanda Brewer

Kami Chauncy

Kristine R. Crews, Pharm.D.

Jun J. Yang, Ph.D.

Alejandro R. Molinelli, Ph.D.

St. Jude Children's Research Hospital®

Finding cures. Saving children.

Thank You

eha Sf(PM)

Dr. Jun J. Yang Lab

Zhenhua Li, Ph.D.
&
Yu-Chih Hsiao, MS

Things I didn't go over that people might be interested in

- What is the magnification of your images? Why did you decide that to be the magnification/FOV# for your assay?
  - Our images are read at 20x, 9-FOVs, & Confocal. Prior testing had been performed before my arrival and these imaging settings were deemed acceptable.
- What Dyes are used for your analysis?
  - AOPI: Control wells
  - CQG: All wells minus AOPI
- When you mention texture, what features are extracted?
  - Mainly SER features such as: Hole, Edge, Ridge, etc.
- Do you use Brightfield channel for any of your data collection blocks?
  - No, we do use this method in other pipelines, and although brightfield can provide a lot of insight on the morphology and texture of a desired cell culture, the MSC co-culture set-up can cause challenges with data collection.
- What type of samples were used for training?
  - PDX, Patient, Pediatric samples
- Would thresholding texture properties make the manual pipeline more accurate?
  - Potentially yes. On the other hand, you are adding a level of complexity for the user to find that goldilocks zone of classification. In return you are increasing data size, increasing analysis time, increasing dependency on thresholding.
- You mentioned your machine learning classifier was a SVM, what is that?
  - A support vector machine (SVM) is a supervised machine learning algorithm that classifies data by finding an optimal line or hyperplane that maximizes the distance between each class in an N-dimensional space.
- How much time is required for a full report on a patient?
  - Generally, 7 days, 4 for incubation/imaging and 3 for quality control and analysis.
- What does SER and STAR mean in your data collection blocks?
  - Spots, Edges, Ridges + Symmetry Threshold, Axial, Radial

https://www.ibm.com/topics/support-vector-machine