

Artificial Intelligence and Disruptive Technologies in Point-of-Care Testing

Center for Diagnostic Innovation



Nam Tran, PhD, HCLD (ABB)
Professor and Director of Clinical Pathology
Director, Center for Diagnostic Innovation

Disclosures

- Consultant for Roche Diagnostics and Roche Molecular Systems
- Consultant for Radiometer
- Received speaking honoraria for Thermo Fisher, DiaSorin, and Nova Biomedical
- Co-inventor of MILO-ML automated ML software, and co-owner of MILO-ML, Inc
- UC Davis Health is a Roche Diagnostic Center of Excellence
- UC Davis is partnered with SpectraPass, LLC to develop a MALDI COVID test

Learning Objectives

- Define artificial intelligence (AI) and machine learning (ML)
- Identify current uses of AI/ML in our daily lives
- Identify AI/ML techniques to detect or prevent testing errors
- Describe AI/ML techniques for image and pattern recognition for diagnostic purposes
- Discuss potential applications for AI generated synthetic data
- Discuss potential uses of AI “chatbots” for POCT applications

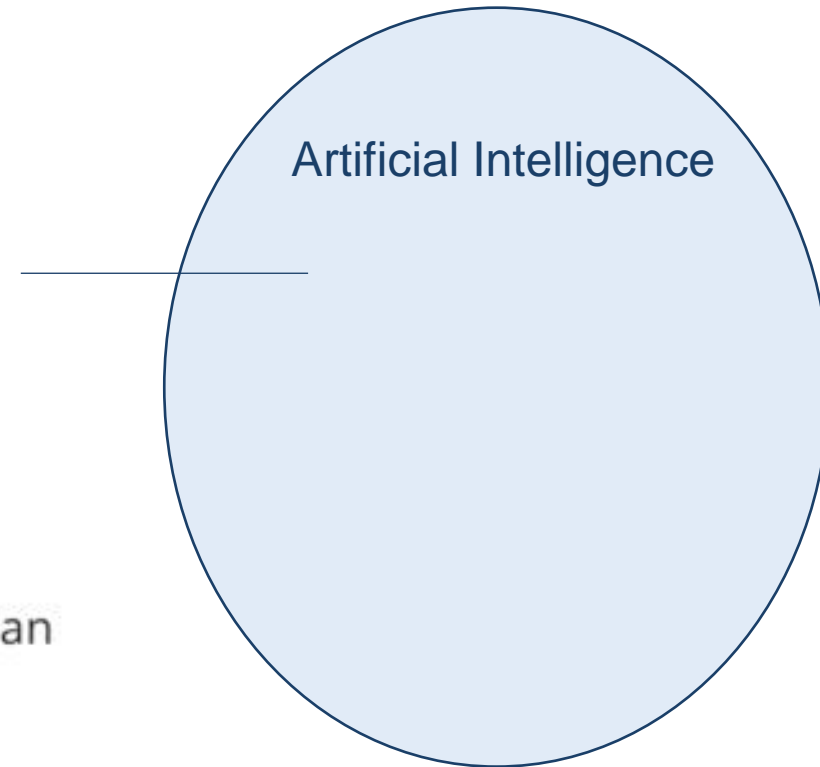
What is Artificial Intelligence / Machine Learning?

artificial intelligence

noun

Definition of *artificial intelligence*

- 1 : a branch of computer science dealing with the simulation of intelligent behavior in computers
- 2 : the capability of a machine to imitate intelligent human behavior



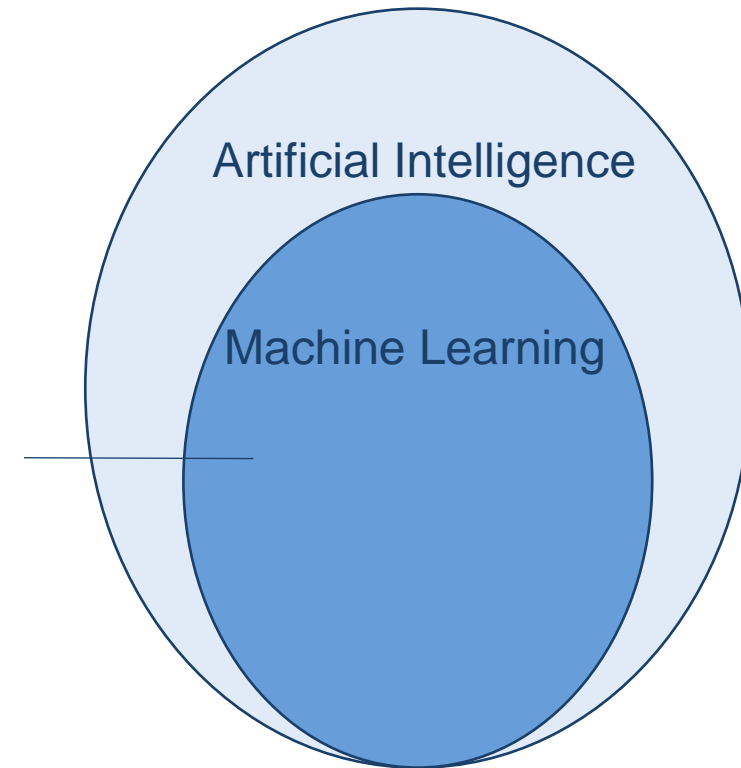
What is Artificial Intelligence / Machine Learning?

machine learning noun

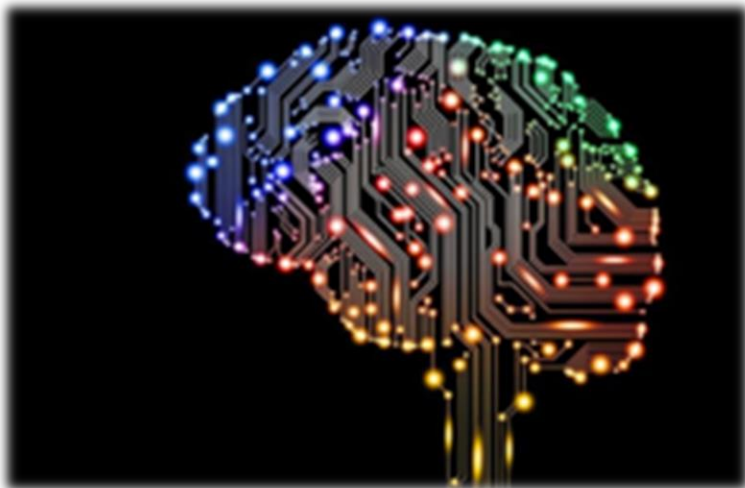
Definition of *machine learning*

: the process by which a computer is able to improve its own performance (as in analyzing image files) by continuously incorporating new data into an existing statistical model

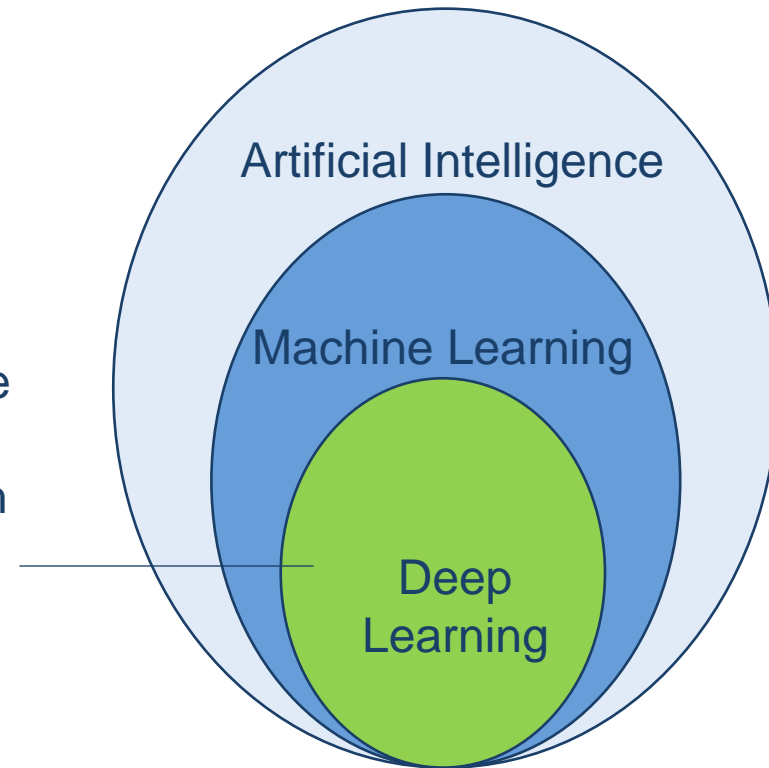
// An entire subspecialty known as *machine learning* is devoted to building algorithms that allow computers to develop new behaviors based on experience.



What is Artificial Intelligence / Machine Learning?



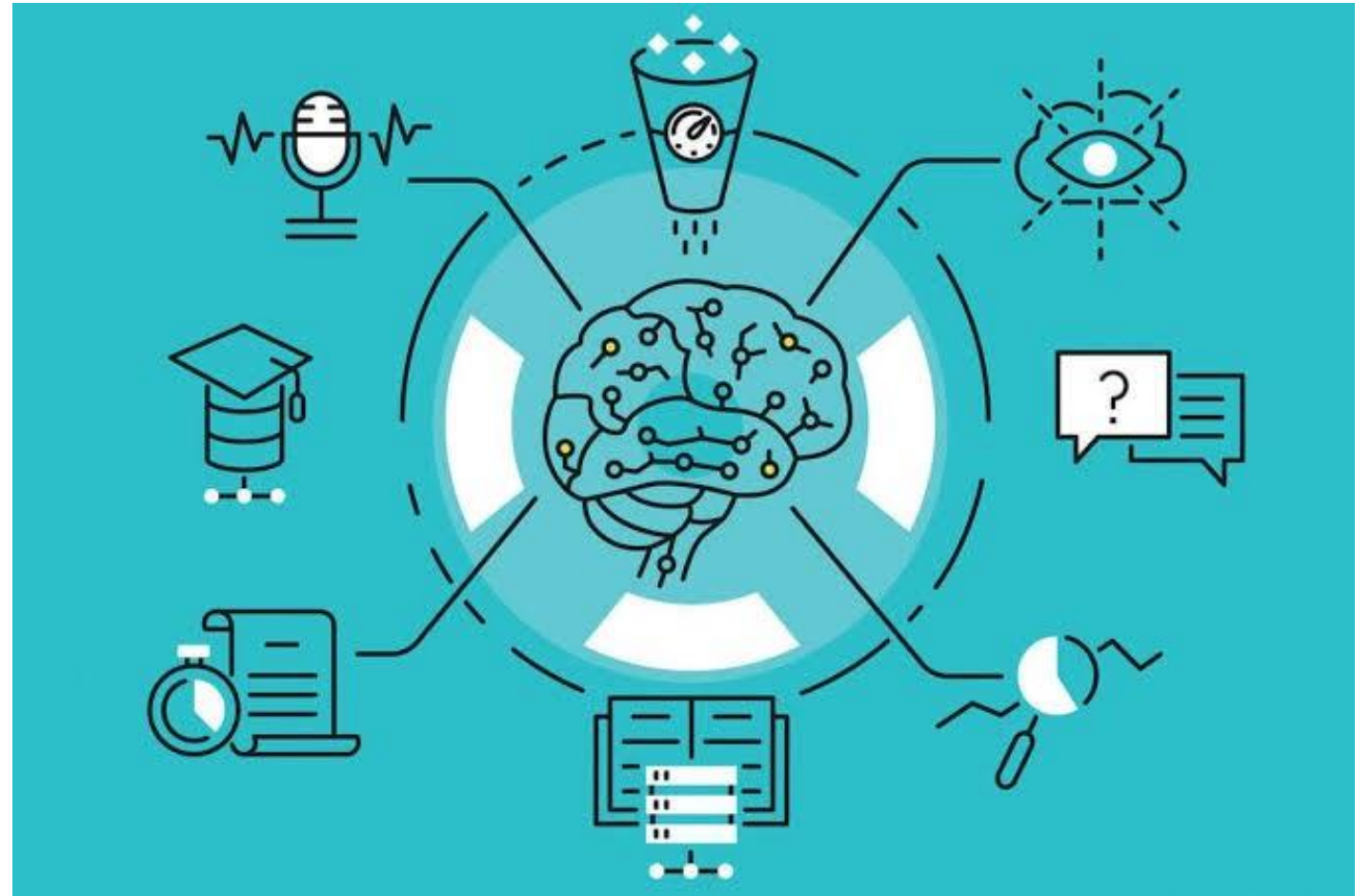
A broader branch of machine learning focused on learning data representations through layers of artificial neural networks.



What is Artificial Intelligence / Machine Learning?

Generative Artificial Intelligence

Generative AI is a type of AI capable of generating text, images or other media in response to user prompts. In short, it uses AI to learn patterns from training data to generate new data that has similar characteristics.



AI/ML is Already Here and its Changing Our Lives!

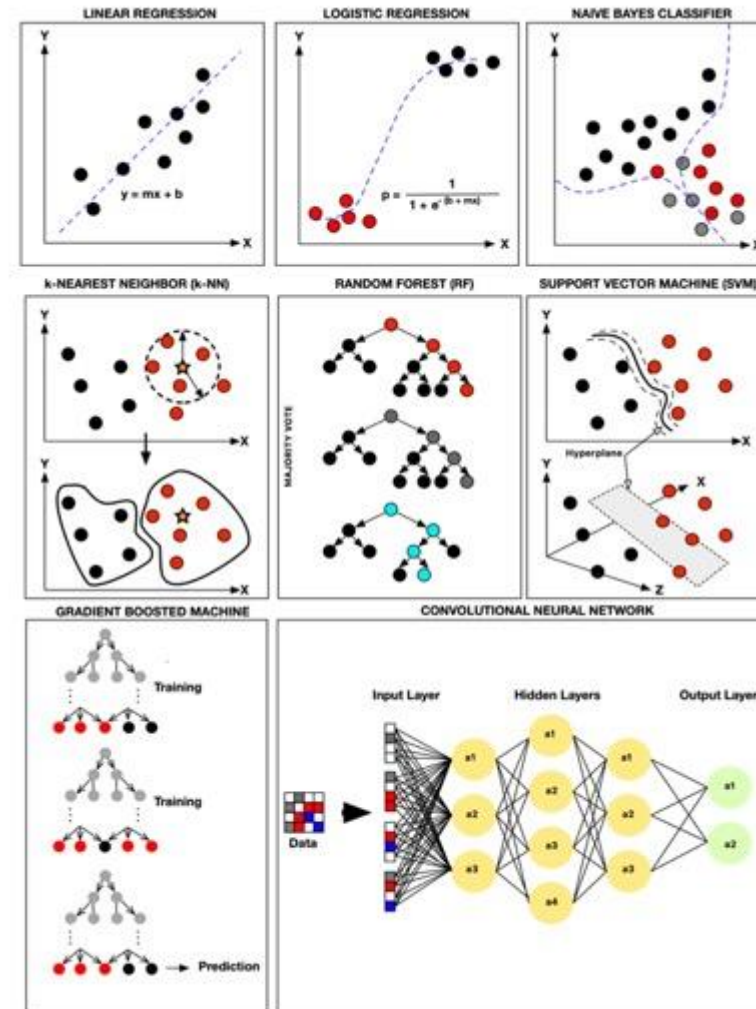


Common AI/ML Methods in Lab Medicine

Parametric

Non-Parametric

- Linear Regression
- Logistic Regression
- Naïve Bayes Classifier
- k-nearest Neighbors
- Random Forest
- Support Vector Machine
- Gradient Boosted Machine
- Convolutional Neural Network



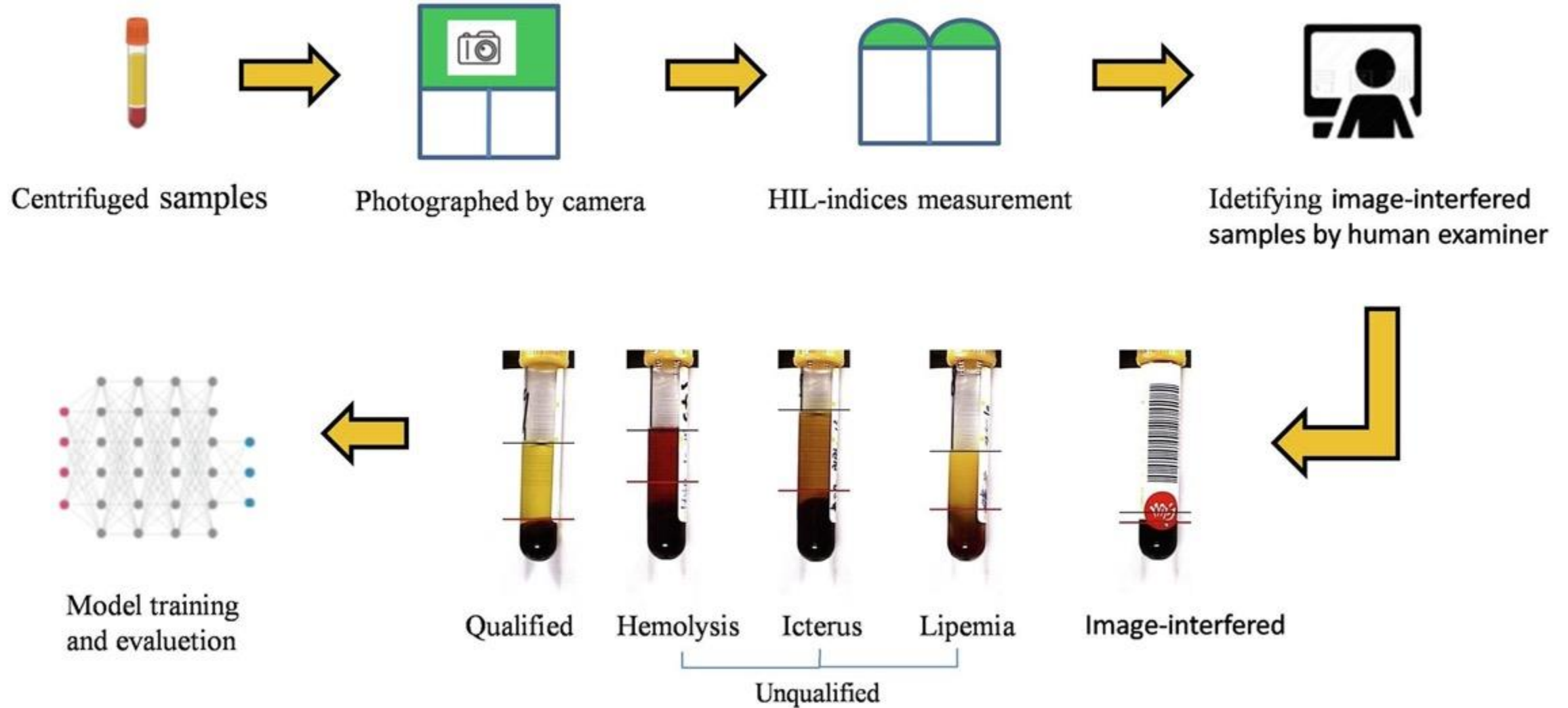
AI/ML Laboratory Opportunities

OPPORTUNITY	EXAMPLES
Error Detection / Prevention	Specimen integrity, mislabeled specimens, facial recognition
Image / Pattern recognition	Slide and fluid analysis
Multi-Analyte / Complex Data Analysis	Mass spectrometry, “big data” applications (e.g., EHR data, genomics, etc)
Automated medical decision-making	Point-of-care testing

AI/ML Laboratory Opportunities

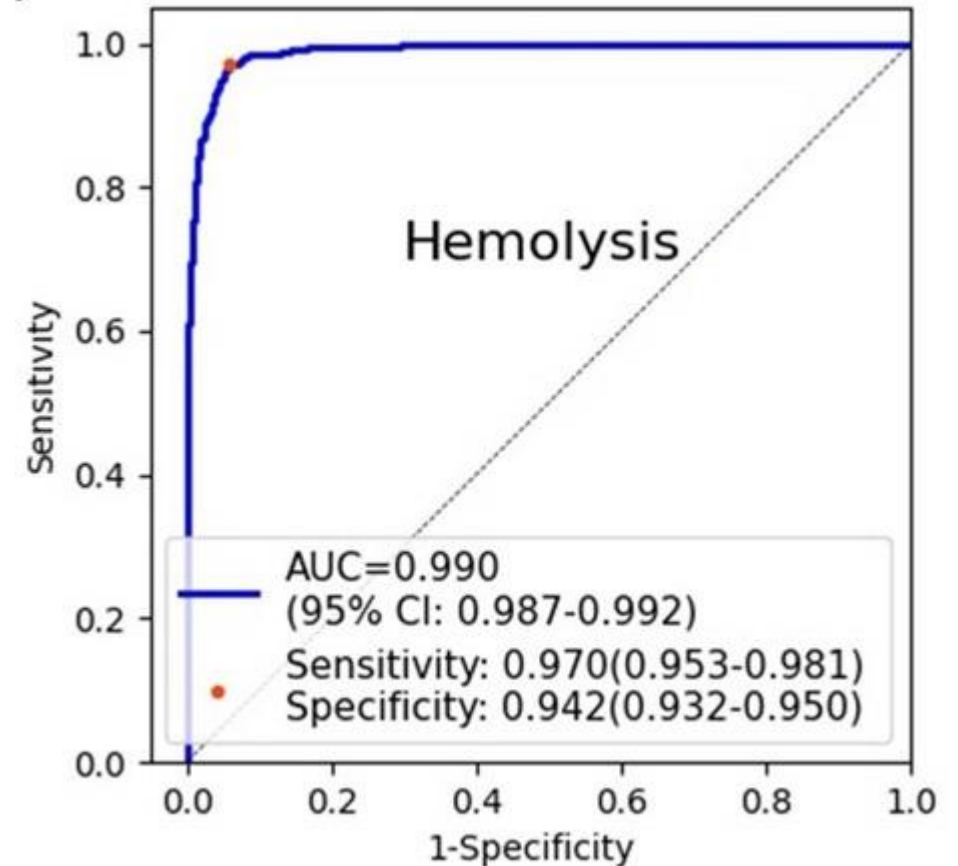
OPPORTUNITY	EXAMPLES
Error Detection / Prevention	Specimen integrity, mislabeled specimens, facial recognition
Image / Pattern recognition	Slide and fluid analysis
Multi-Analyte / Complex Data Analysis	Mass spectrometry, “big data” applications (e.g., EHR data, genomics, etc)
Automated medical decision-making	Point-of-care testing

Hemolysis Detection by Machine Learning



Hemolysis Detection by Machine Learning

- A dataset of 16,427 sample images were collected. Samples divided into training and test datasets.
- Deep learning algorithms achieved an area of the curve 0.99 with a sensitivity of 97% and specificity of 94.2%.
- When the algorithm was applied, serum indices tests were reduced for 26.76% (n = 1225)
- Among the 26.76% of samples, 1 sample was misclassified for predicted serum indices lower than the lower limit but measured serum indices greater than the cutoff values, 3 samples were misclassified for predicted serum indices greater than the upper limit but measured serum indices lower than the cutoff values





Volume 150, Issue 6
December 2018

[< Previous](#)

FEATURED

Using Machine Learning-Based Multianalyte Delta Checks to Detect Wrong Blood in Tube Errors

Matthew W Rosenbaum, MD, Jason M Baron, MD ✉

American Journal of Clinical Pathology, Volume 150, Issue 6, 24 October 2018, Pages 555–566,

<https://doi.org/10.1093/ajcp/aqy085>

Published: 30 August 2018

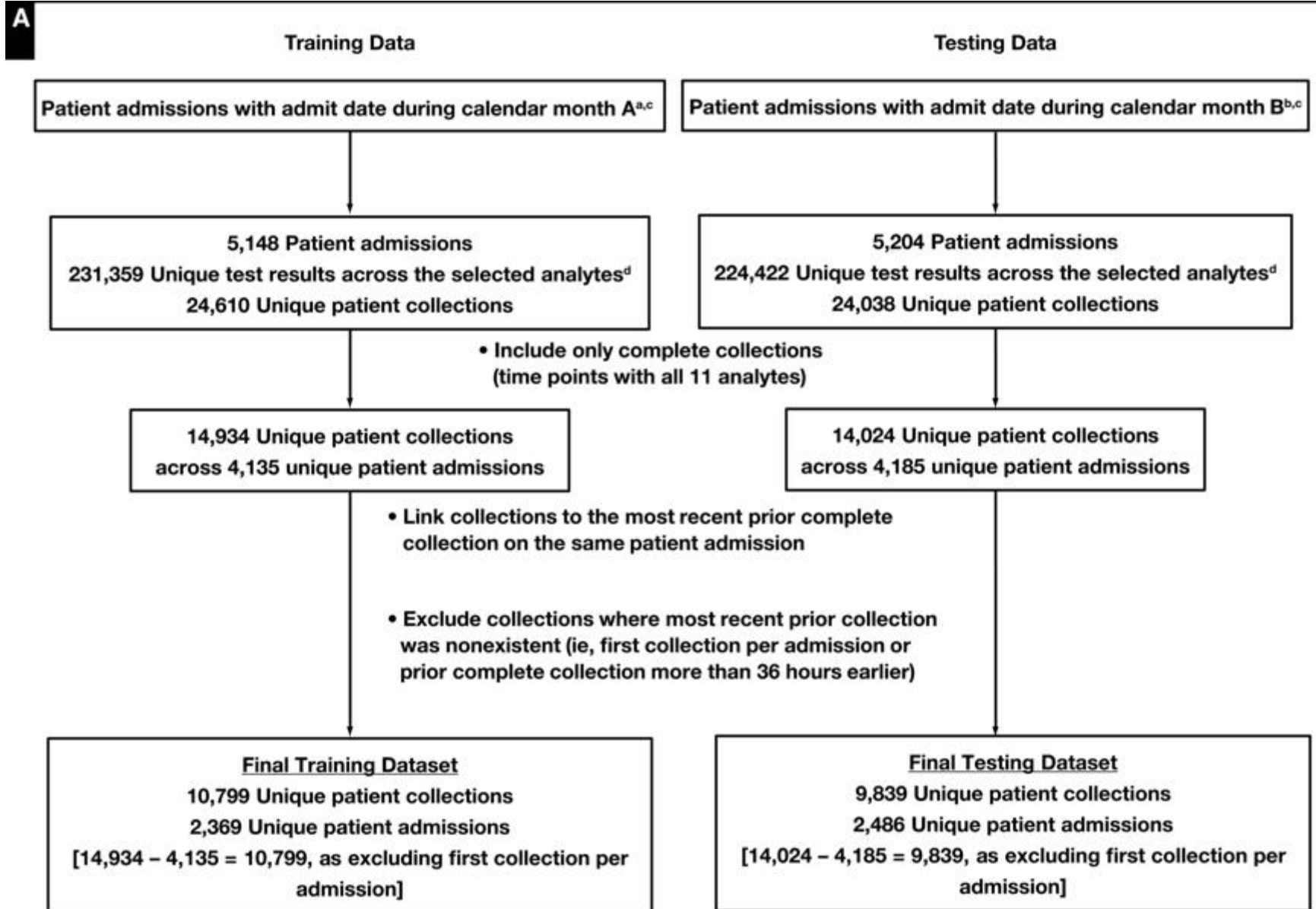
“ Cite 🔑 Permissions ➦ Share ▼

Abstract

Objectives

An unfortunate reality of laboratory medicine is that blood specimens collected from one patient occasionally get mislabeled with identifiers from a different patient, resulting in so-called “wrong blood in tube” (WBIT) errors and potential patient harm. Here, we sought to develop a machine learning-based, multianalyte delta check algorithm to detect WBIT errors and mitigate patient harm.

Study Methods: Overall Design



Study Methods

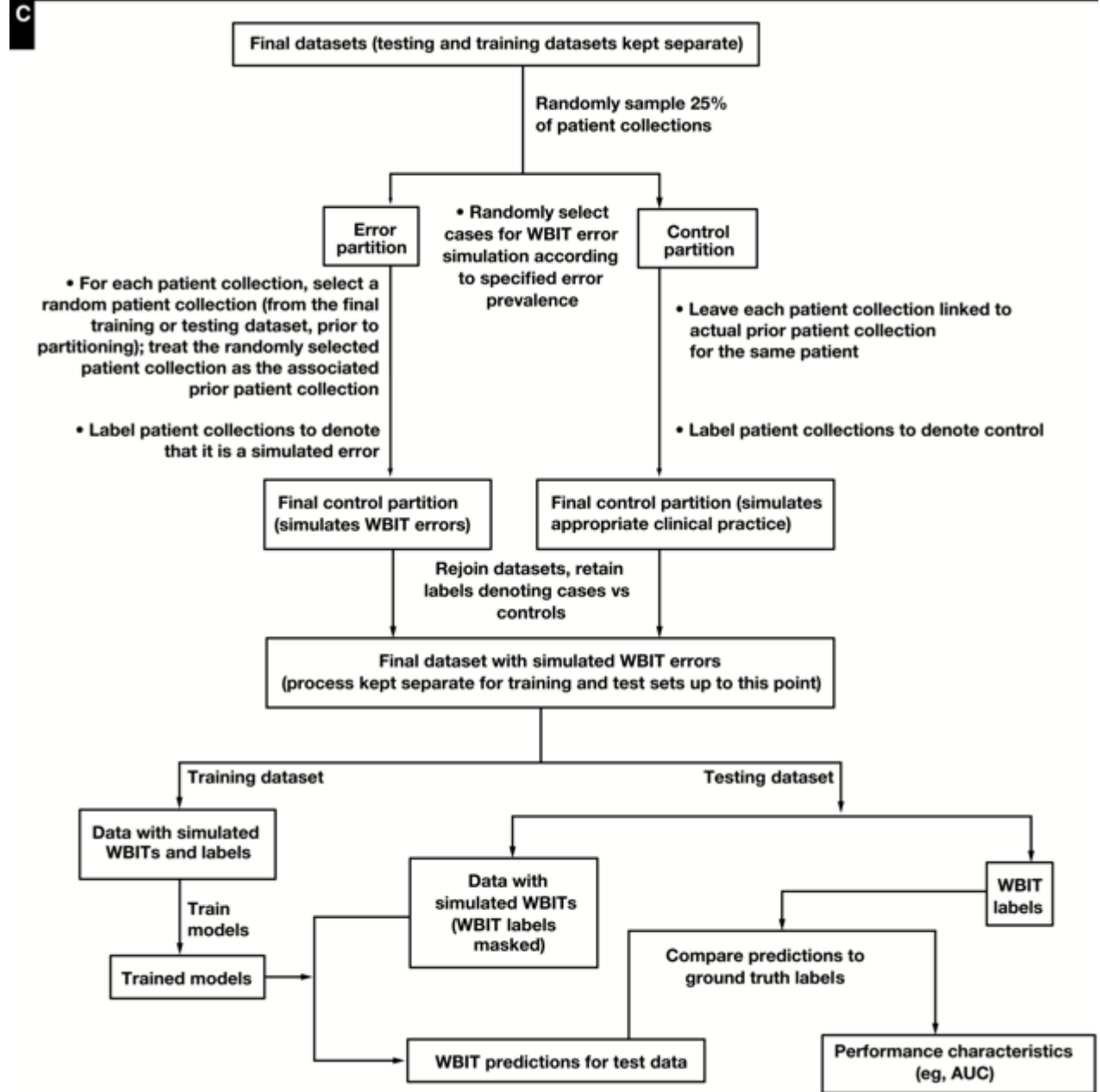
B

Original Data							
Patient admission	Collection date/time	Na	K	...	Prior Na	Prior K	...
1234567 - 1/1/1990	1/2/1990 6 AM	140	3.9	...	--	--	
1234567 - 1/1/1990	1/3/1990 6 AM	141	3.8	...	140	3.9	
2234567 - 1/1/1990	1/2/1990 6 AM	142	3.6	...	--	--	
2234567 - 1/1/1990	1/3/1990 6 AM	143	3.7	...	142	3.6	
3234567 - 1/1/1990	1/2/1990 6 AM	131	5.1	...	--	--	
3234567 - 1/1/1990	1/3/1990 6 AM	133	5.0	...	131	5.1	



After WBIT Error Simulation								
Patient admission	Collection date/time	Na	K	...	Prior Na	Prior K	...	Case/control
1234567 - 1/1/1990	1/2/1990 6 AM	140	3.9	...	--	--		Excluded, no prior results
1234567 - 1/1/1990	1/3/1990 6 AM	141	3.8	...	140	3.9		Control
2234567 - 1/1/1990	1/2/1990 6 AM	142	3.6	...	--	--		Excluded, no prior results
2234567 - 1/1/1990	1/3/1990 6 AM	133	5.0	...	142	3.6		WBIT error case
3234567 - 1/1/1990	1/2/1990 6 AM	131	5.1	...	--	--		Excluded, no prior results
3234567 - 1/1/1990	1/3/1990 6 AM	133	5.0	...	131	5.1		Control

Patient 3234567 had a specimen mislabeled with a label from patient 2234567

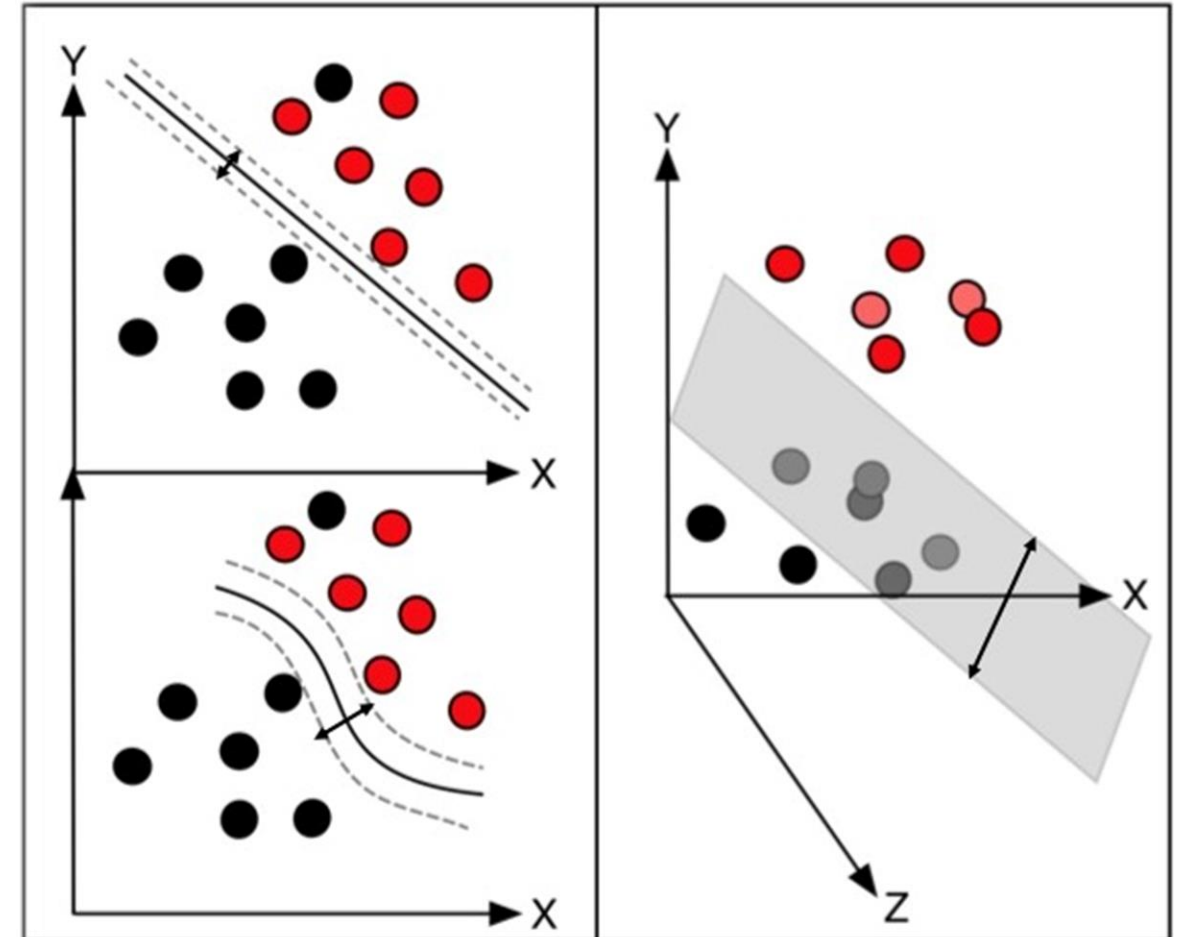


Methods of Analysis including AI/ML Techniques

Model Name	Type	Predictors
Univariate models		
Univariate absolute difference (named for each analyte)	Univariate: evaluate sensitivity/specific at various thresholds	Absolute change in consecutive results for each analyte
Univariate velocity	Univariate: evaluate sensitivity/specific at various thresholds	Absolute velocity of change between consecutive results for each analyte
Multivariate models		
Logistic regression, difference only	Logistic regression	Absolute change in consecutive results for each analyte
Logistic regression, velocity only	Logistic regression	Absolute velocity of change between consecutive results for each analyte
Logistic regression, difference and values	Logistic regression	(1) Absolute change in consecutive results for each analyte; (2) actual test results
SVM, difference only	SVM	Absolute change in consecutive results for each analyte
SVM, difference and values	SVM	(1) Absolute change in consecutive results for each analyte; (2) actual test results

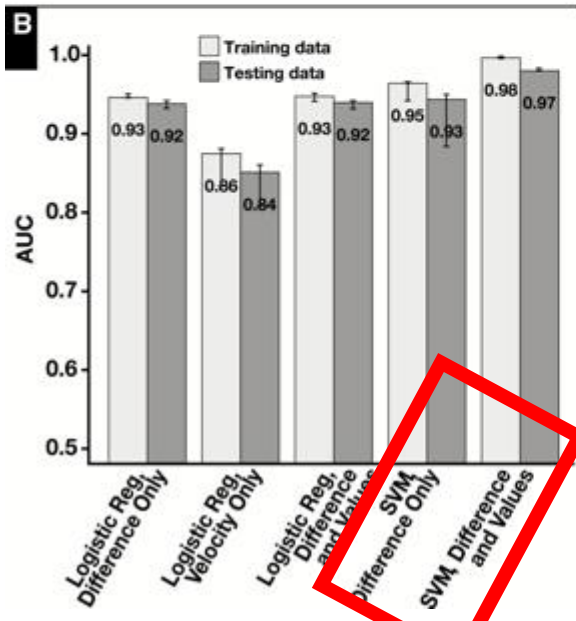
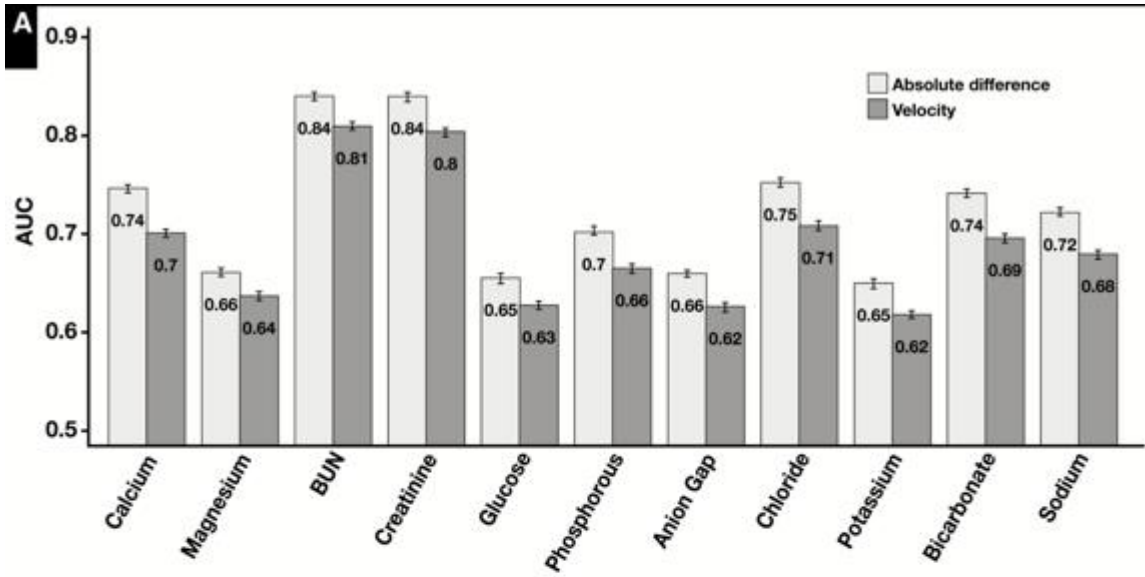
SVM, support vector machines.

What is Support Vector Machine (SVM)

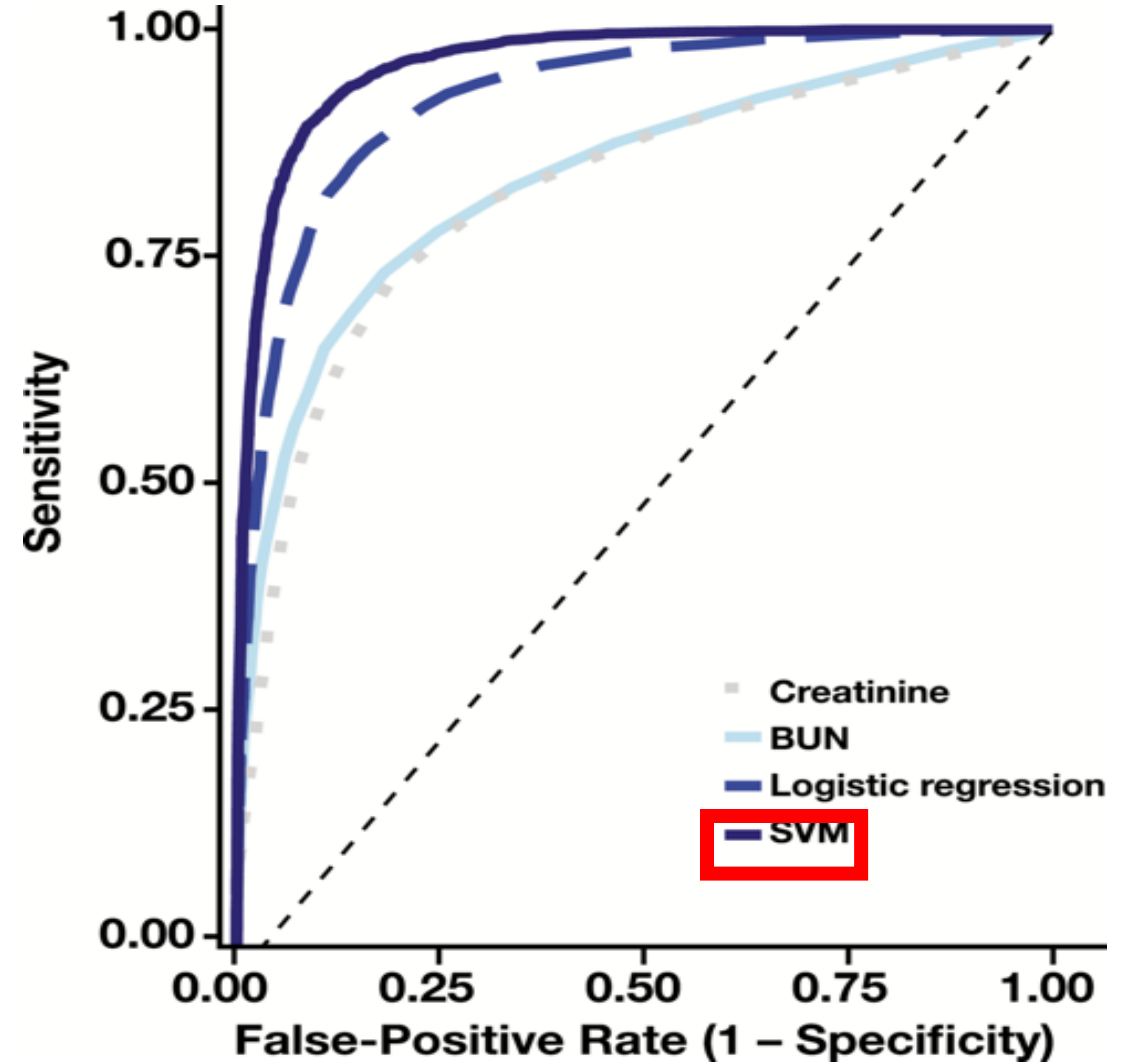


- Constructs a hyperplane (—) that best separates groups.
- The best hyperplane maximizing the margins (---) is selected.
- Hyperplanes may exist in 3D space to improve separation of data points and further maximize margins.

Results – Predictive Power of AI/ML (SVM) for WBIT Events



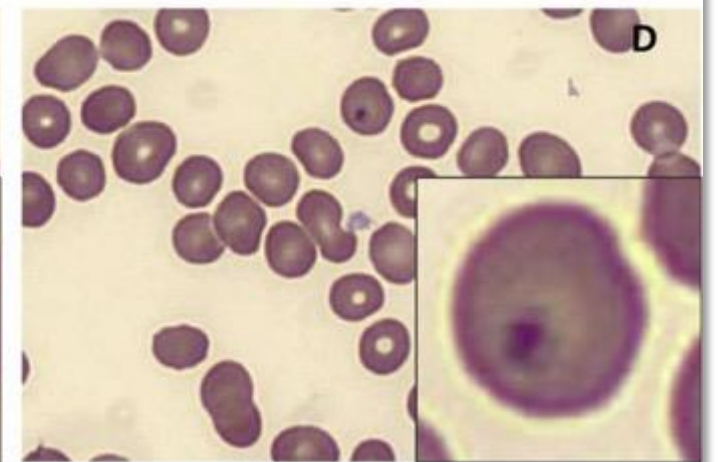
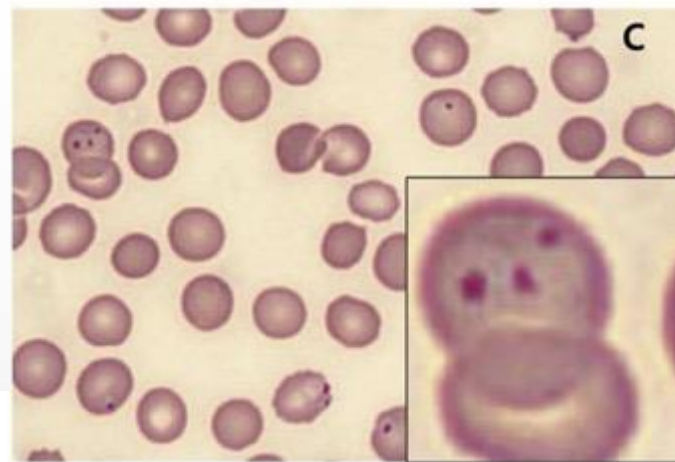
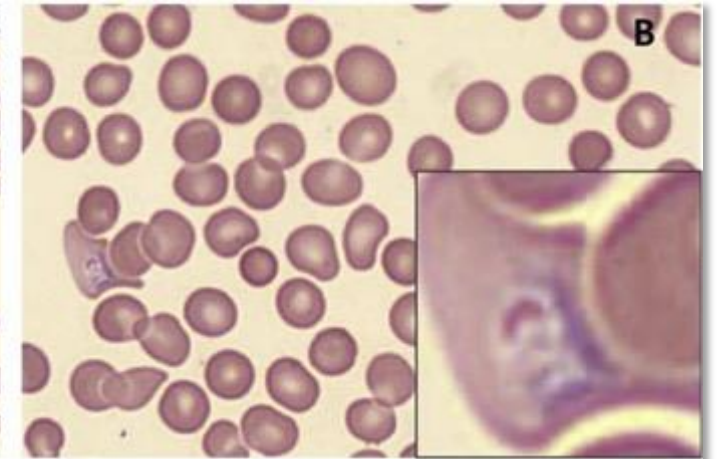
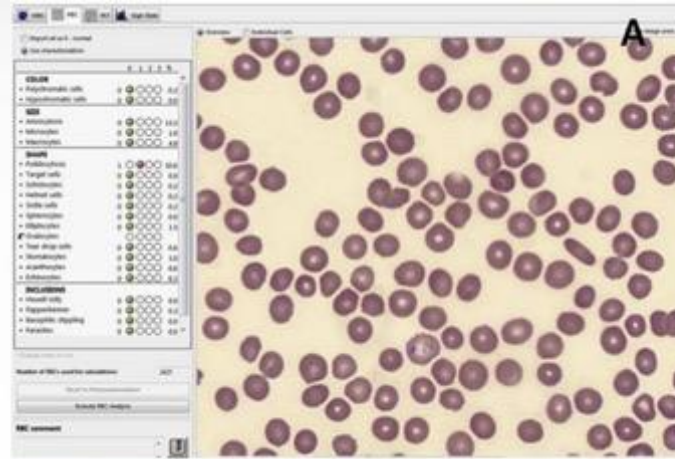
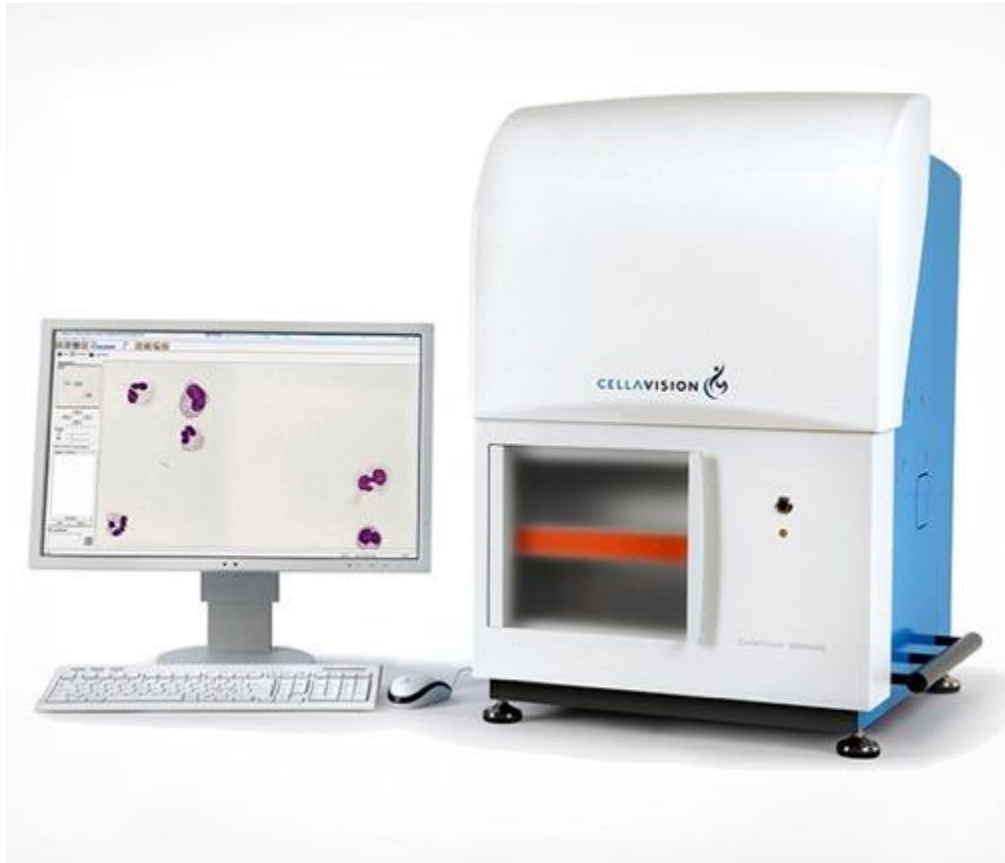
SVM performed better than other traditional statistical methods such as logistic regression when evaluating lab value differences alone and/or with values.



AI/ML Laboratory Opportunities

OPPORTUNITY	EXAMPLES
Error Detection / Prevention	Specimen integrity, mislabeled specimens, facial recognition
Image / Pattern recognition	Slide and fluid analysis
Multi-Analyte / Complex Data Analysis	Mass spectrometry, “big data” applications (e.g., EHR data, genomics, etc)
Automated medical decision-making	Point-of-care testing

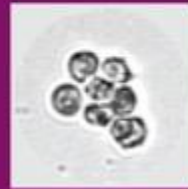
Hematology Image Recognition



Automated Urinalysis Image Recognition



White Blood Cells



White Blood Cell Clumps



Red Blood Cells



Squamous Epithelial Cells



Non-Squamous Epititheial Cells



Unclassified Casts



Hyaline Casts



Unclassified Crystals



Bacteria



Yeast



Mucus

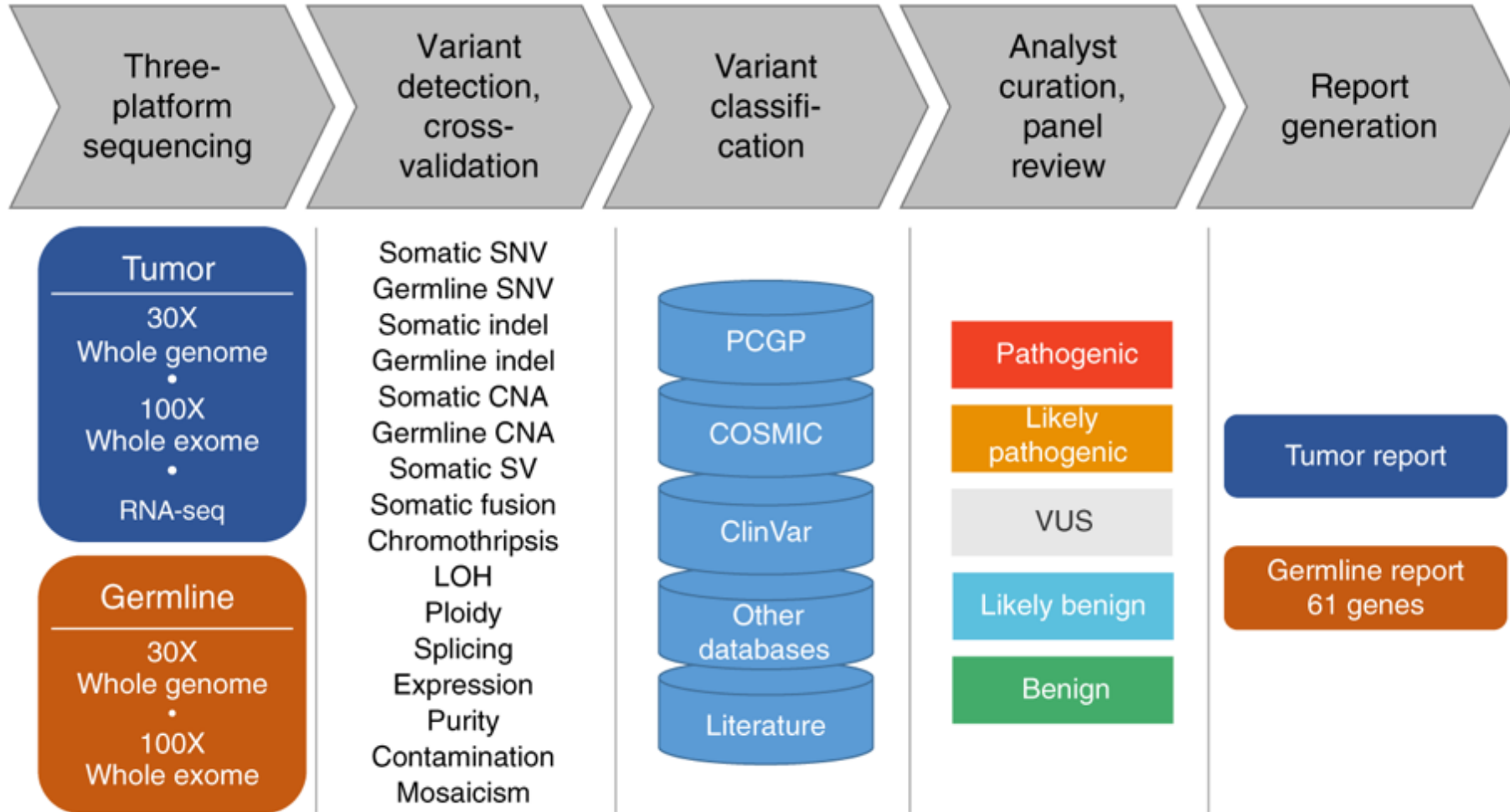


Sperm

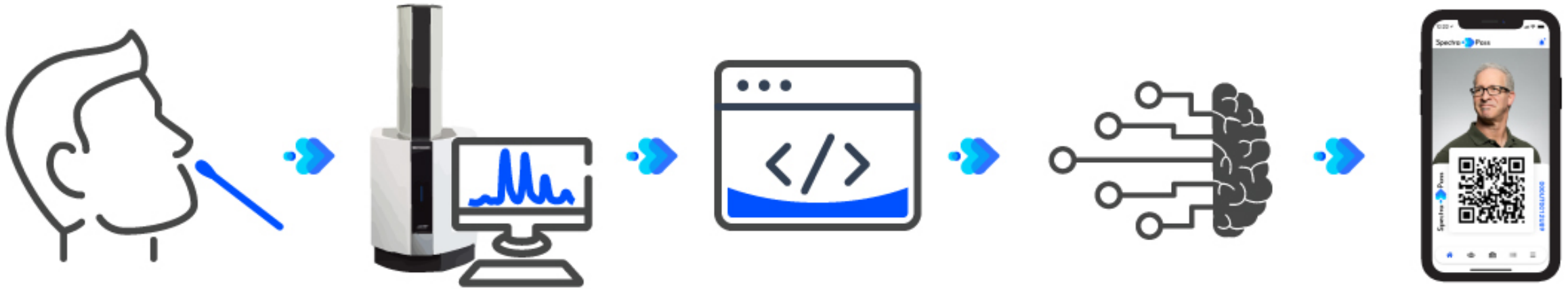
AI/ML Laboratory Opportunities

OPPORTUNITY	EXAMPLES
Error Detection / Prevention	Specimen integrity, mislabeled specimens, facial recognition
Image / Pattern recognition	Slide and fluid analysis
Multi-Analyte / Complex Data Analysis	Mass spectrometry, “big data” applications (e.g., EHR data, genomics, etc)
Automated medical decision-making	Point-of-care testing

Artificial Intelligence for Genomic Testing



AI/ML Enhanced Mass Spec COVID-19 Testing Solution



Step 1

- Clinical specimen acquisition
- Method translation to clinical testing
- PCR reference method

Step 2

- MALDI-TOF Generated MS peaks

Step 3

- MS peaks standardization software

Step 4

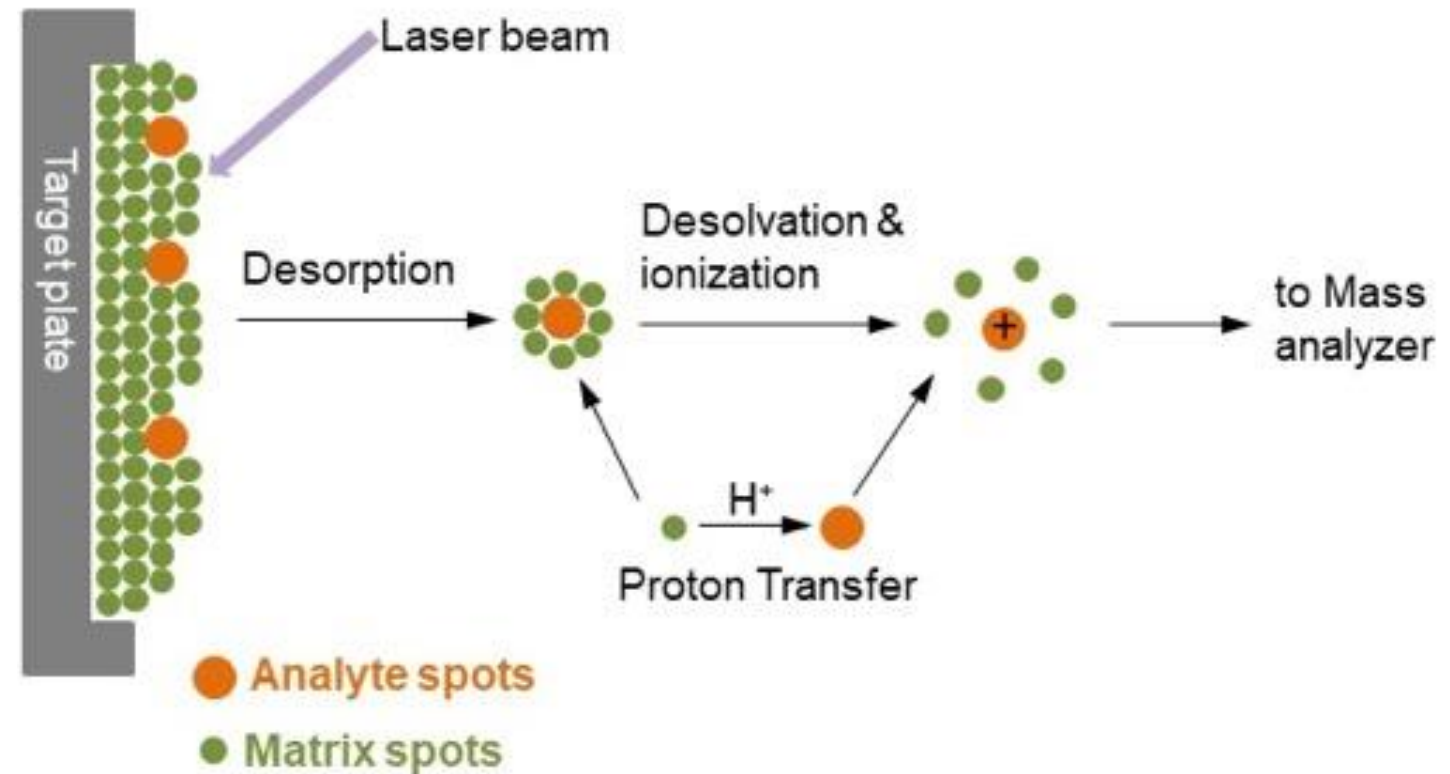
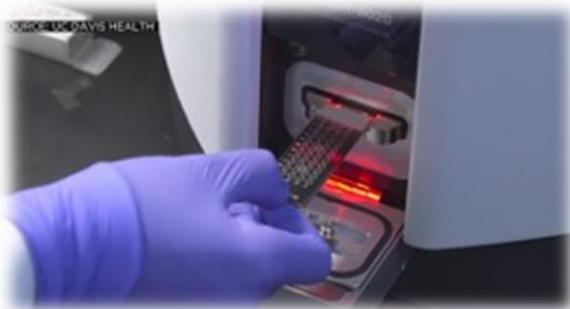
- MILO (Machine Intelligence Learning Optimizer) Auto-ML generating thousands of ML models
- Best MILO ML model selected

Step 5

- New MALDI Mass Spec test analyzed through ML-based prediction on new cases
- Communicating with Spectra-Pass App

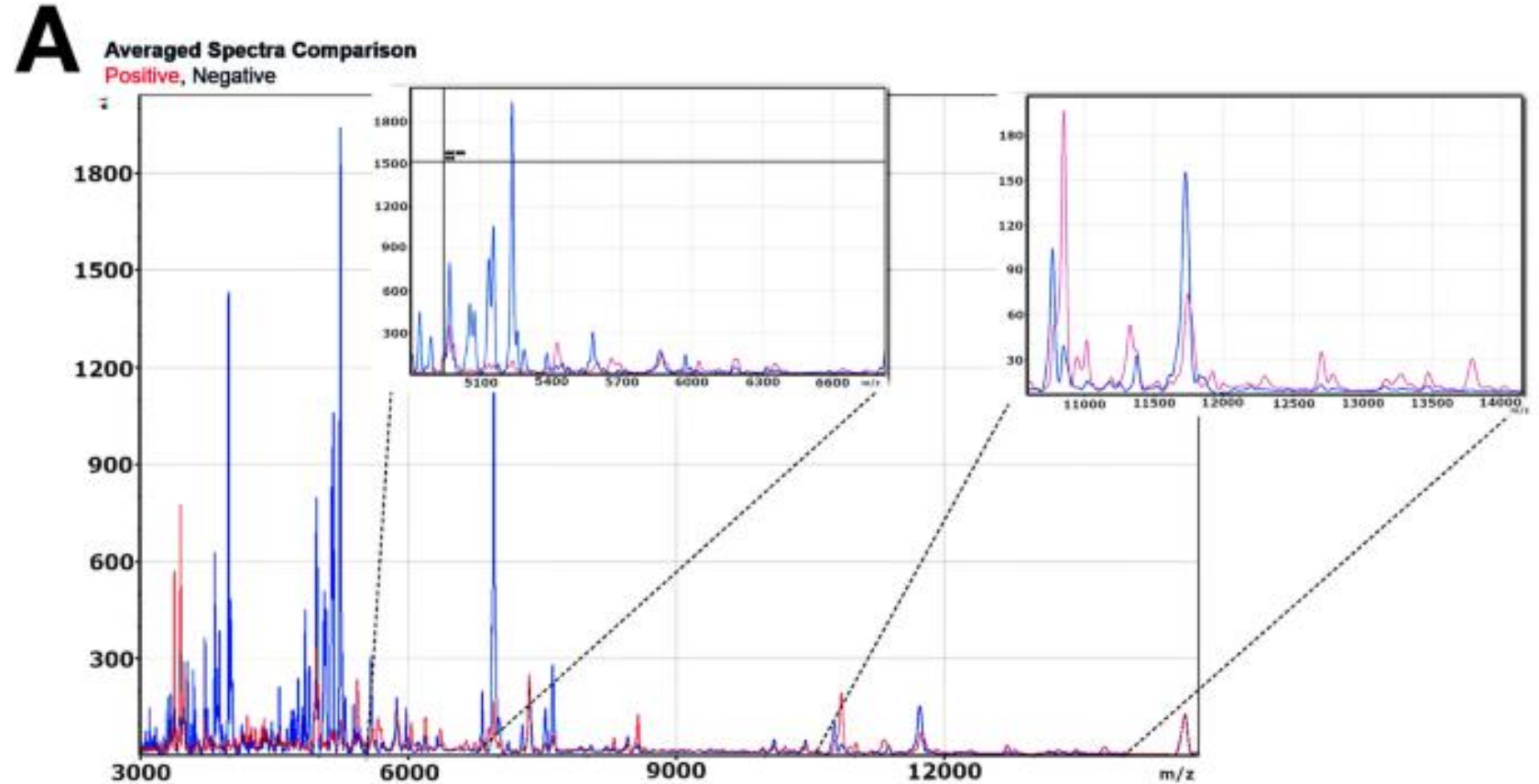
- As a way to bypass COVID-19 molecular and antigen testing supply shortages, UCDH and partners with SpectraPass developed a novel mass spectrometry-based COVID-19 test that also leveraged machine learning.
- Machine learning as based on the in-house developed Machine Intelligence Learning Optimizer (MILO) platform.

MALDI-TOF-MS Based COVID-19 Testing Solution

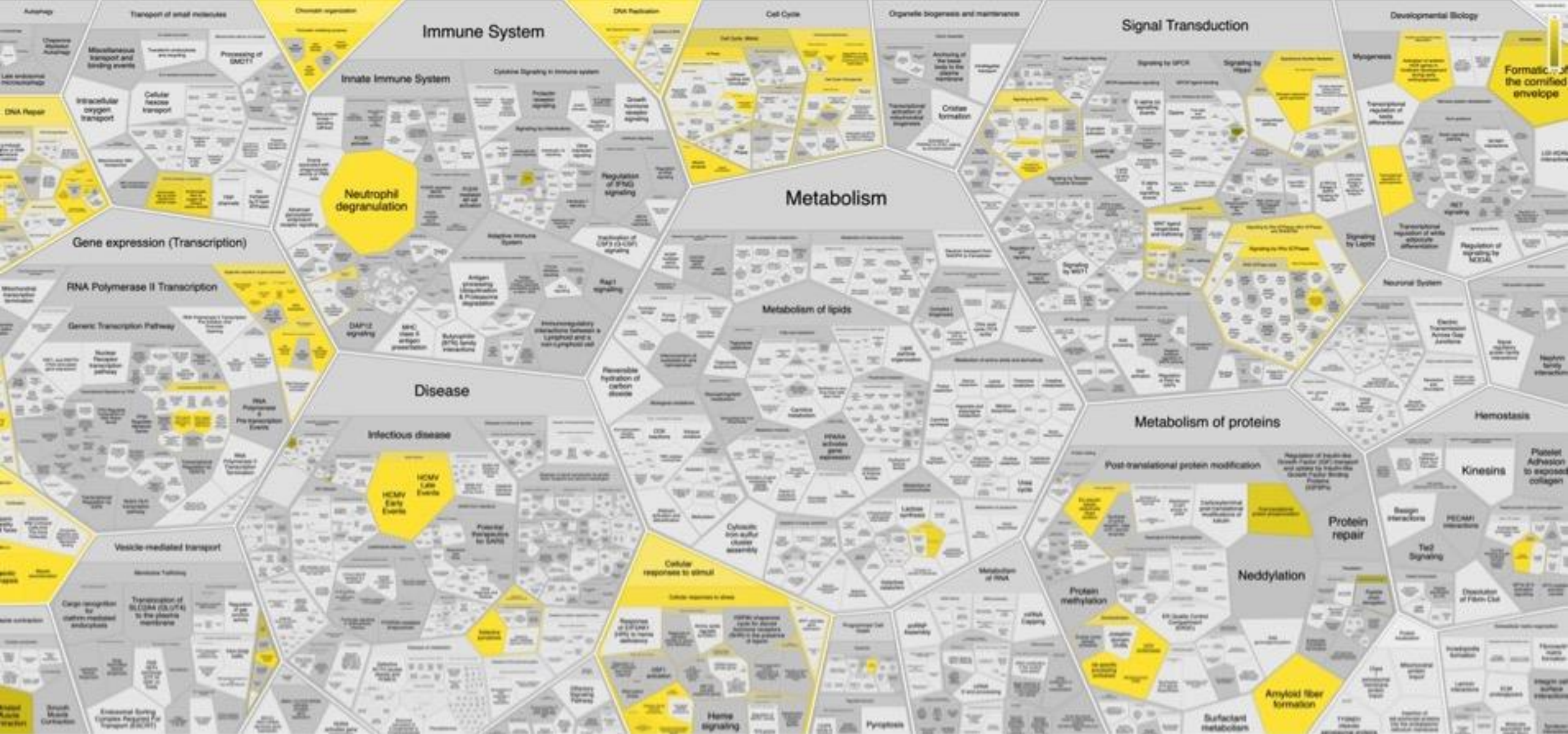


MALDI-TOF-MS provides means to detect host-response profiles (proteins) from clinical samples. Low cost, fast, high-throughput, and can be random access.

Machine Learning Enhanced MALDI-TOF-MS-Based Detection of COVID-19



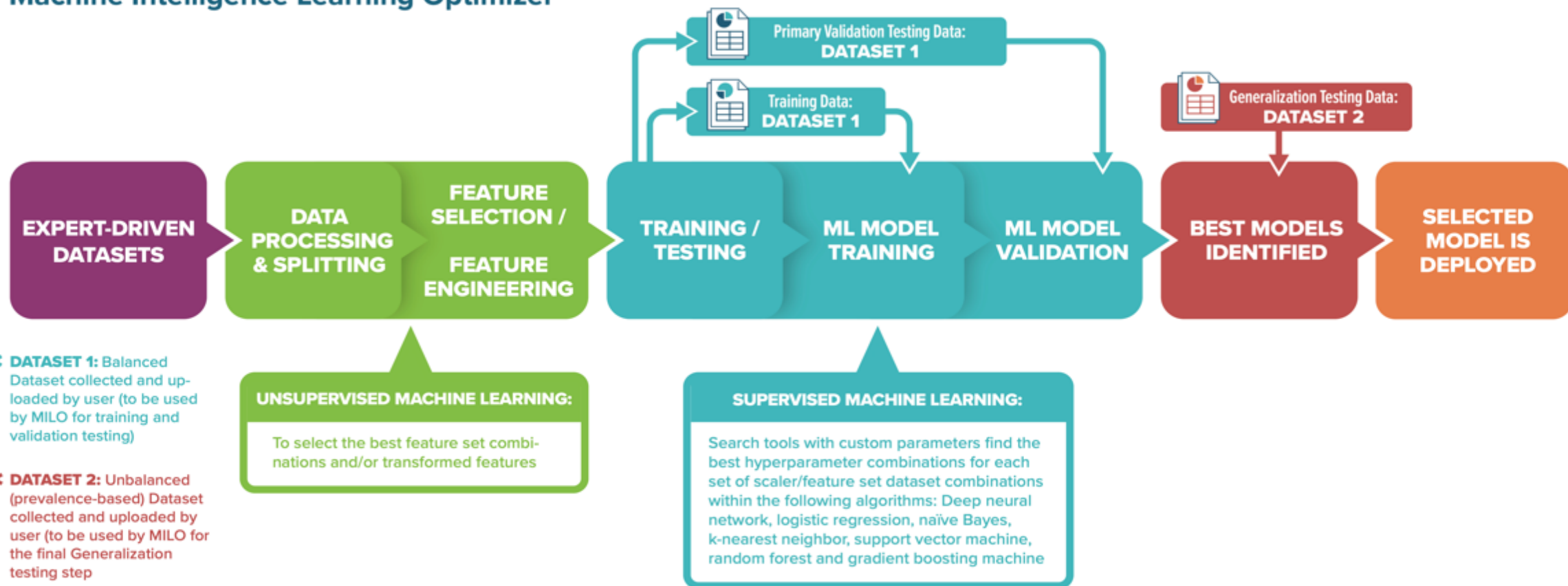
Mass spectrometry is a unique way to analyze samples of various kinds. For COVID-19, anterior nares proteins detected by MALDI-TOF-MS produces hundreds of peaks. How do you analyze this data??



Our MALDI study also identified 14,270 endogenous peptides across 1,198 protein groups. This "Reactome" figure highlights protein association with certain pathways. Yellow highlights indicate high association. Darker shades of gray indicate poor (non-statistically significant) pathway association (Tsai H, *et al. ACS Omega* 2022;20:17462)

MILO AUTO-ML PLATFORM

Machine Intelligence Learning Optimizer

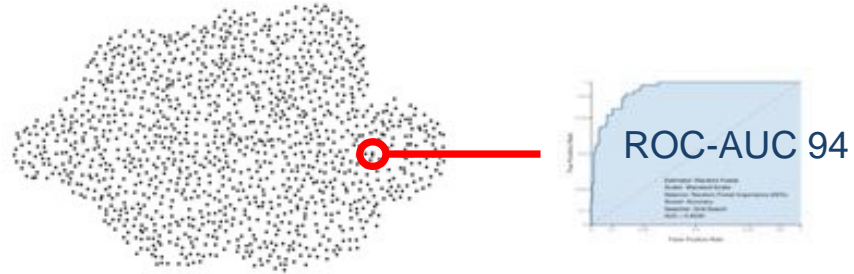


NOTES

- MILO removes missing values in Datasets automatically
- Dataset's number and percentages within each category is automatically assessed
- Data is scaled per pipeline using of variety of scaling methods
- MILO automatically builds and deploys the user-selected model in a WebApp (this live model can now be used to test new data and make predictions)
- The PMML or joblib file can also be easily exported as per user's needs

MORE MODELS – LESS TIME – MORE OPPORTUNITIES

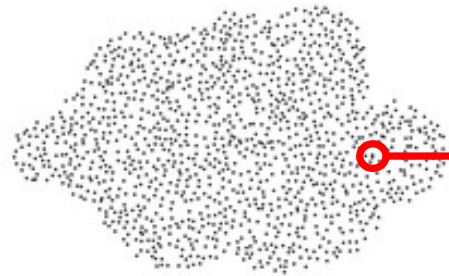
Current ML Approach



49,940 models
~400 hours (~4 months)

MORE MODELS – LESS TIME – MORE OPPORTUNITIES

Current ML Approach



49,940 models
~400 hours (~4 months)



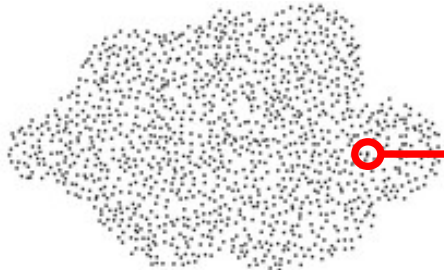
?

?

Are there any other potential models?

MORE MODELS – LESS TIME – MORE OPPORTUNITIES

Current ML Approach



49,940 models
~400 hours (~4 months)

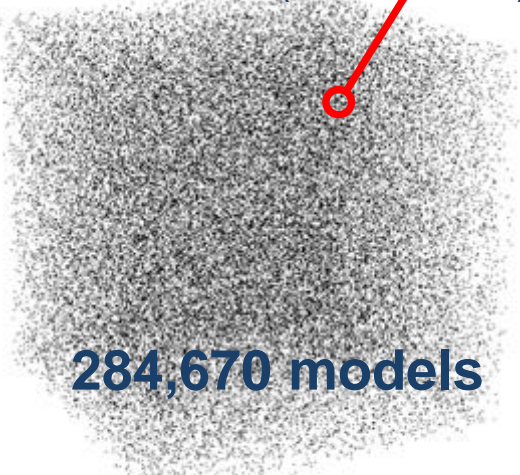


?

?

Are there any other potential models?

MILO

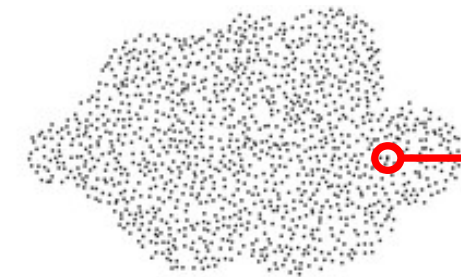


284,670 models
<24 hours

MILO found the model identified via manual programming

MORE MODELS – LESS TIME – MORE OPPORTUNITIES

Current ML Approach



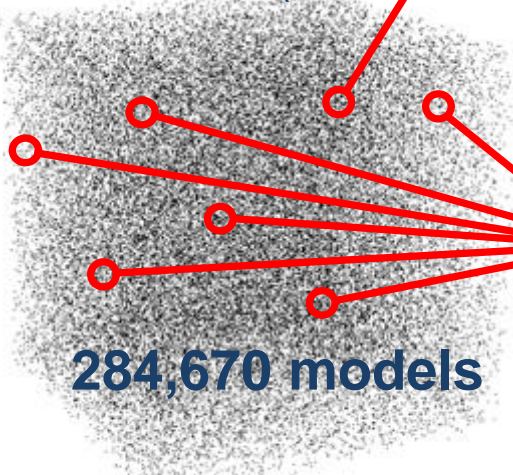
49,940 models
~400 hours (~4 months)



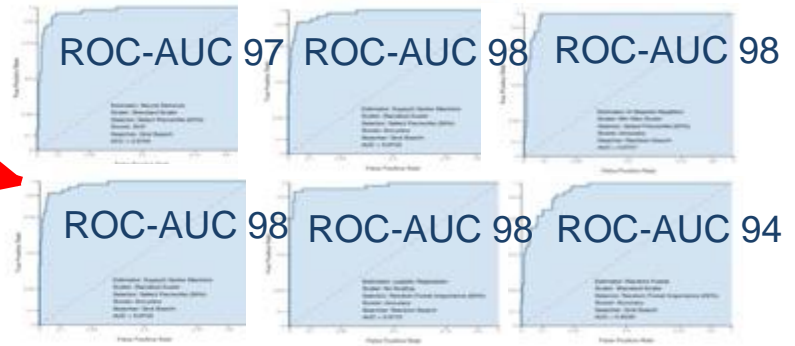
? ?

Are there any other potential models?

MILO



284,670 models
<24 hours



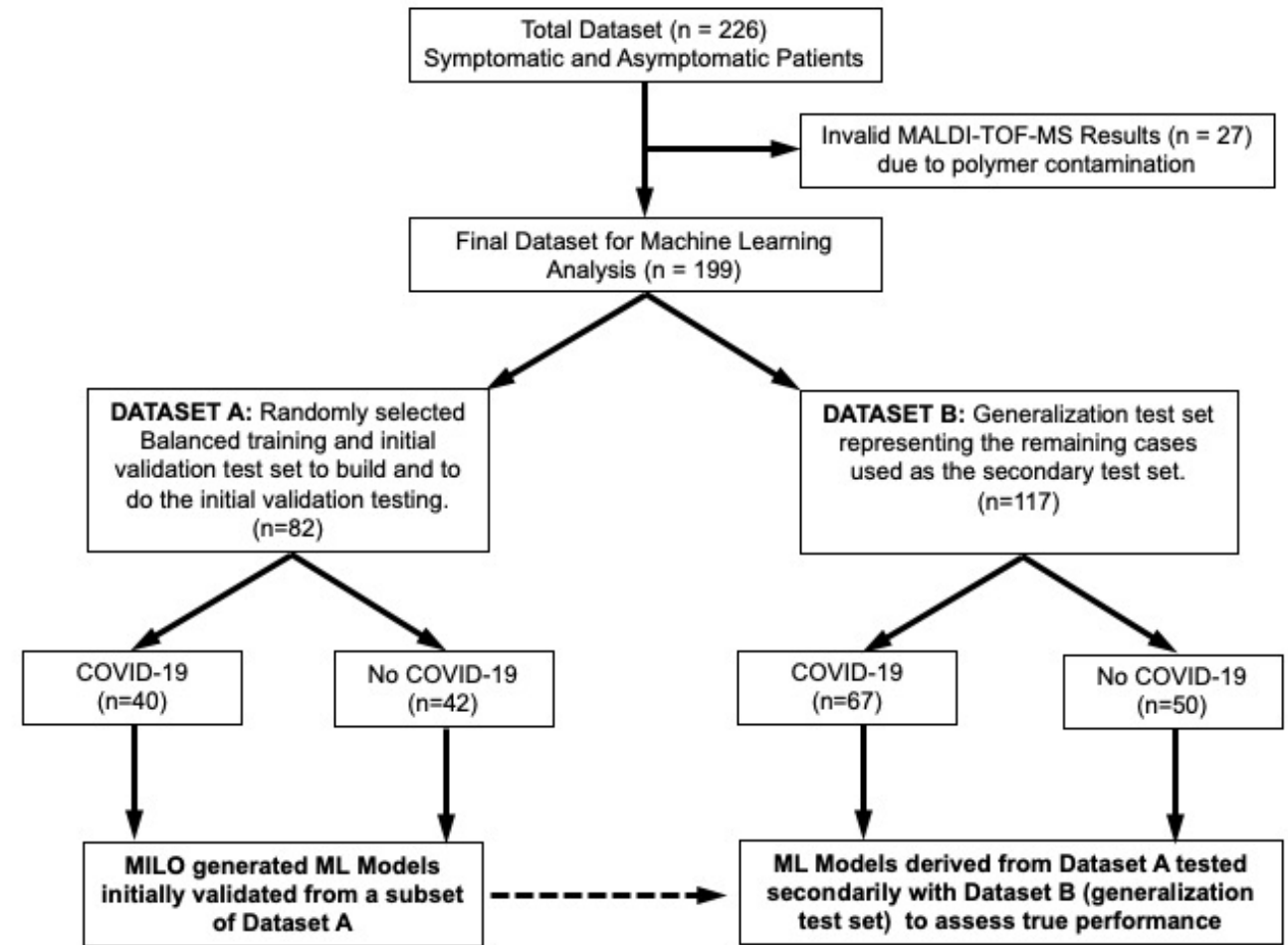
MILO found six additional better models not found by the Current Traditional ML approach

Limitations of manual programming

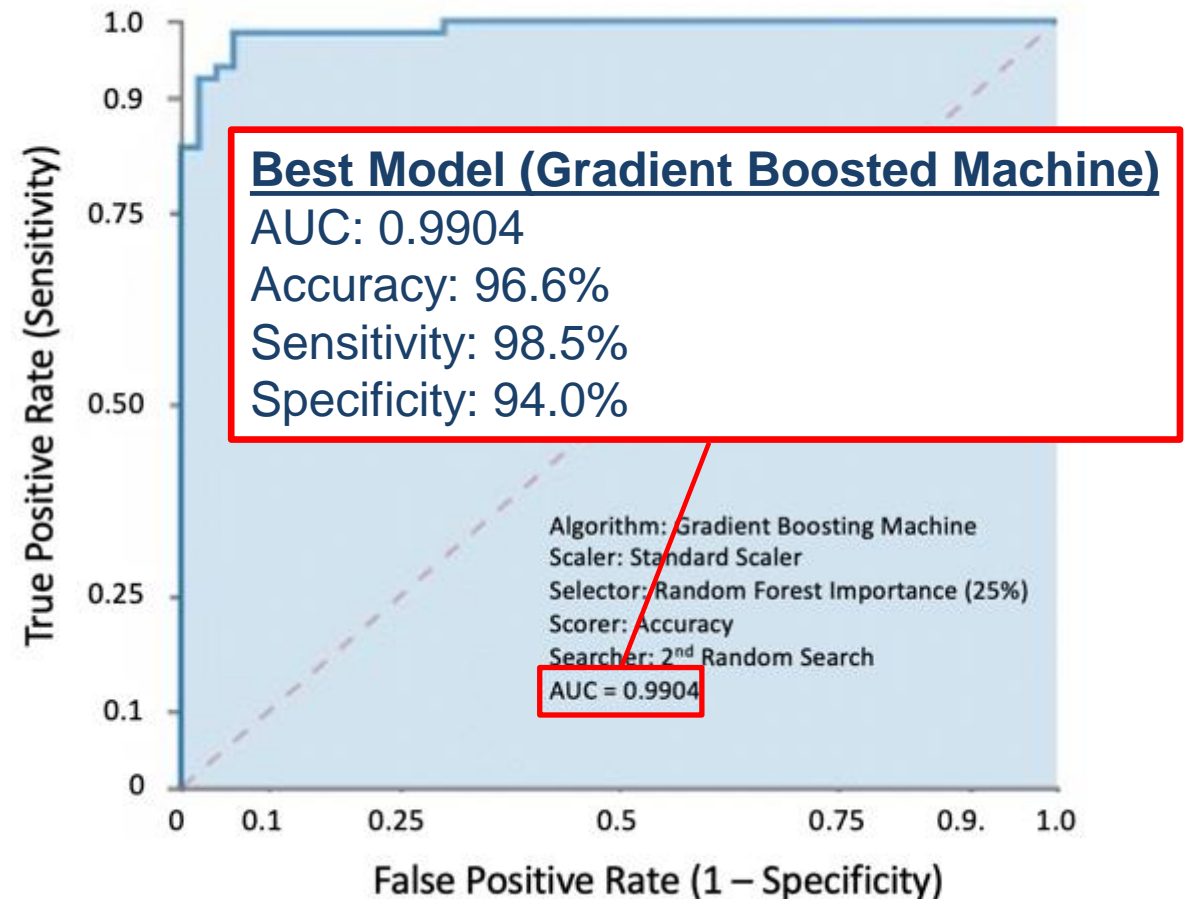
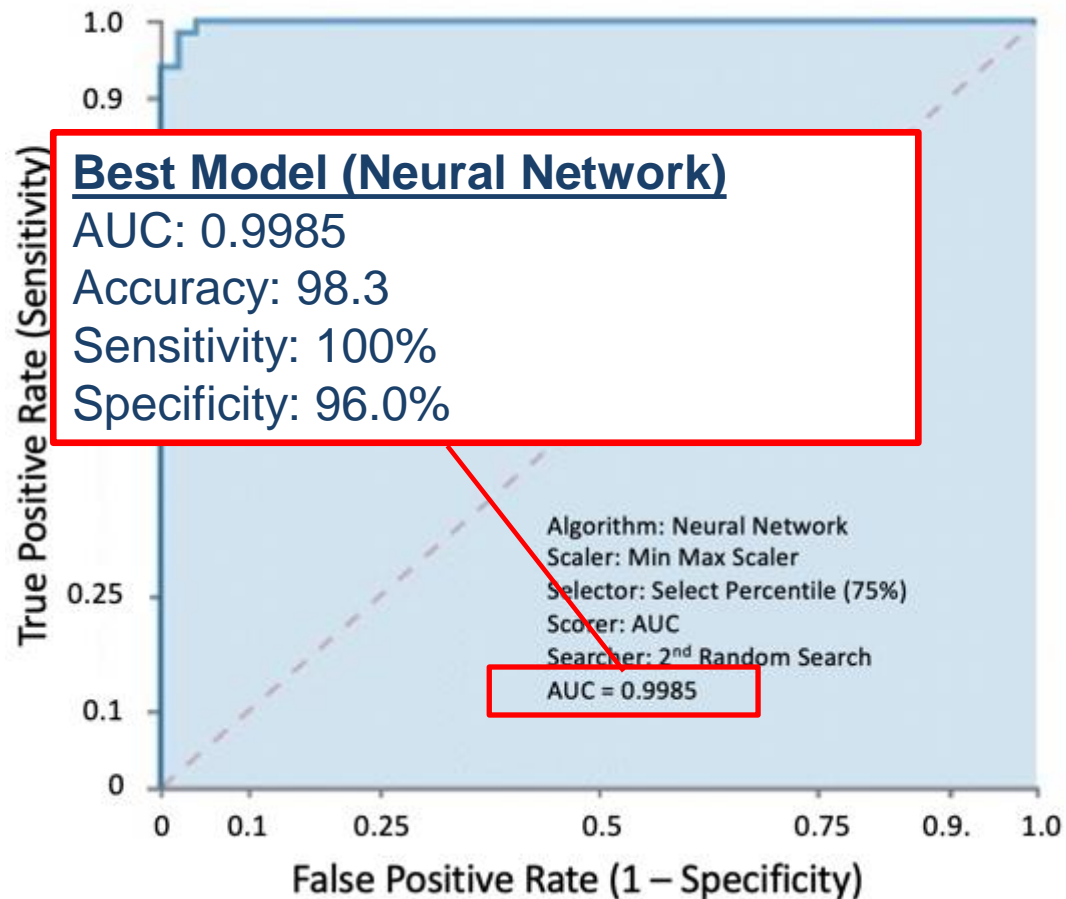
- Time
- Number of combinations
- Bias towards "preferred methods"
- Bias towards feature combinations

Machine Learning Enhanced MALDI-TOF-MS-Based Detection of COVID-19: Pilot Performance

- Pilot study was based on both archived and prospectively collected anterior nares swab specimens
- Conducted from Aug to Dec 2020
- NGS performed to identify presence of non-COVID-19 pathogens, as well as type COVID-19 variants.
- Goal of the study was to confirm we can readily differentiate between COVID-19 from normal patients, including symptomatic vs. asymptomatic cases.

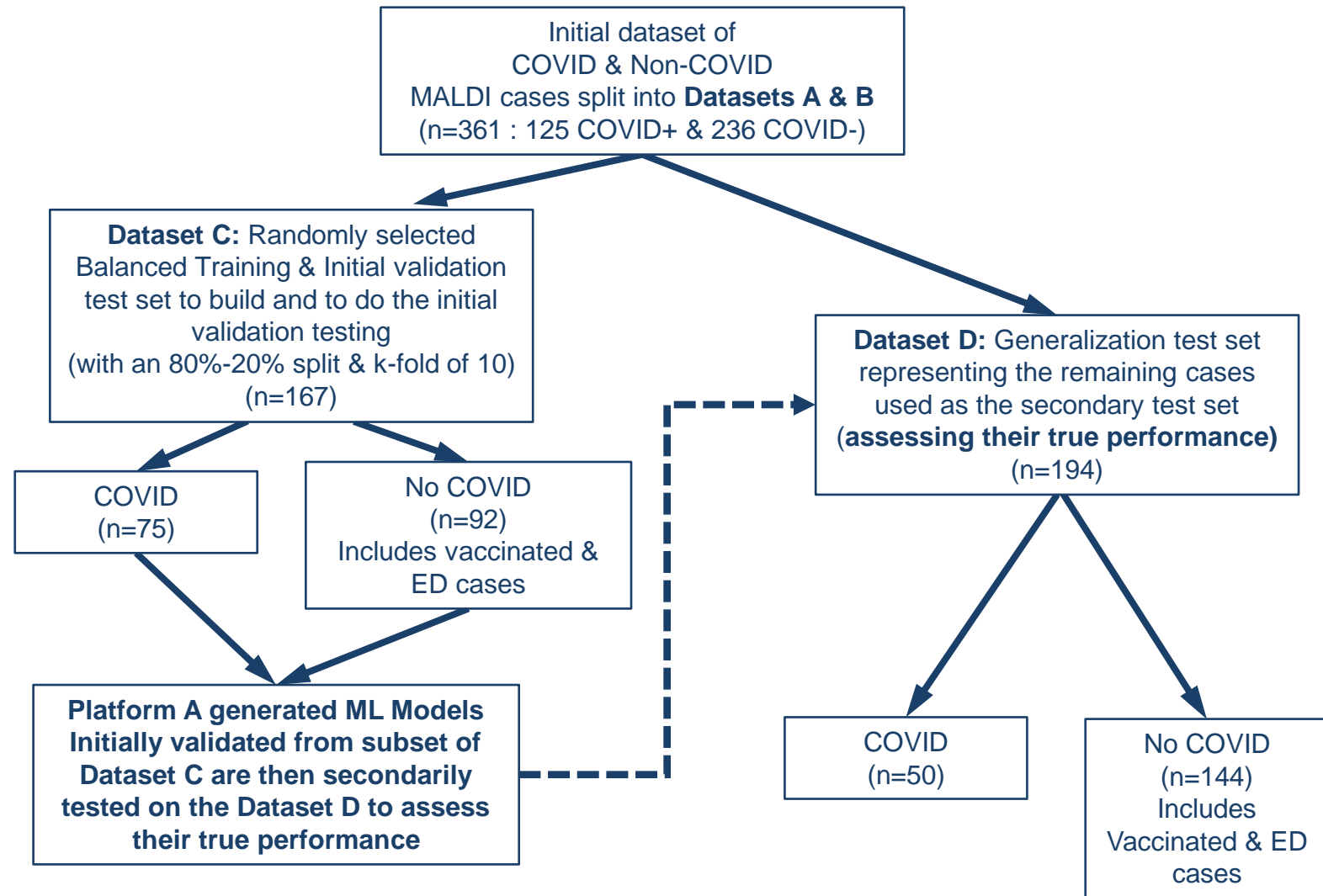


Machine Learning Enhanced MALDI-TOF-MS-Based Detection of COVID-19: Pilot Performance



Proving to Ourselves this is Real – Further Generalization and Validation!

- A second study was started in early 2021 using the original 199 subjects, plus a more heterogeneous group to total 361 patients.
- Added a second site (Las Vegas)
- Included vaccinated individuals
- Detected some non-COVID-19 pathogens by NGS (n = 3) and several SARS-CoV-2 variants (alpha, gamma, delta, iota, and others)



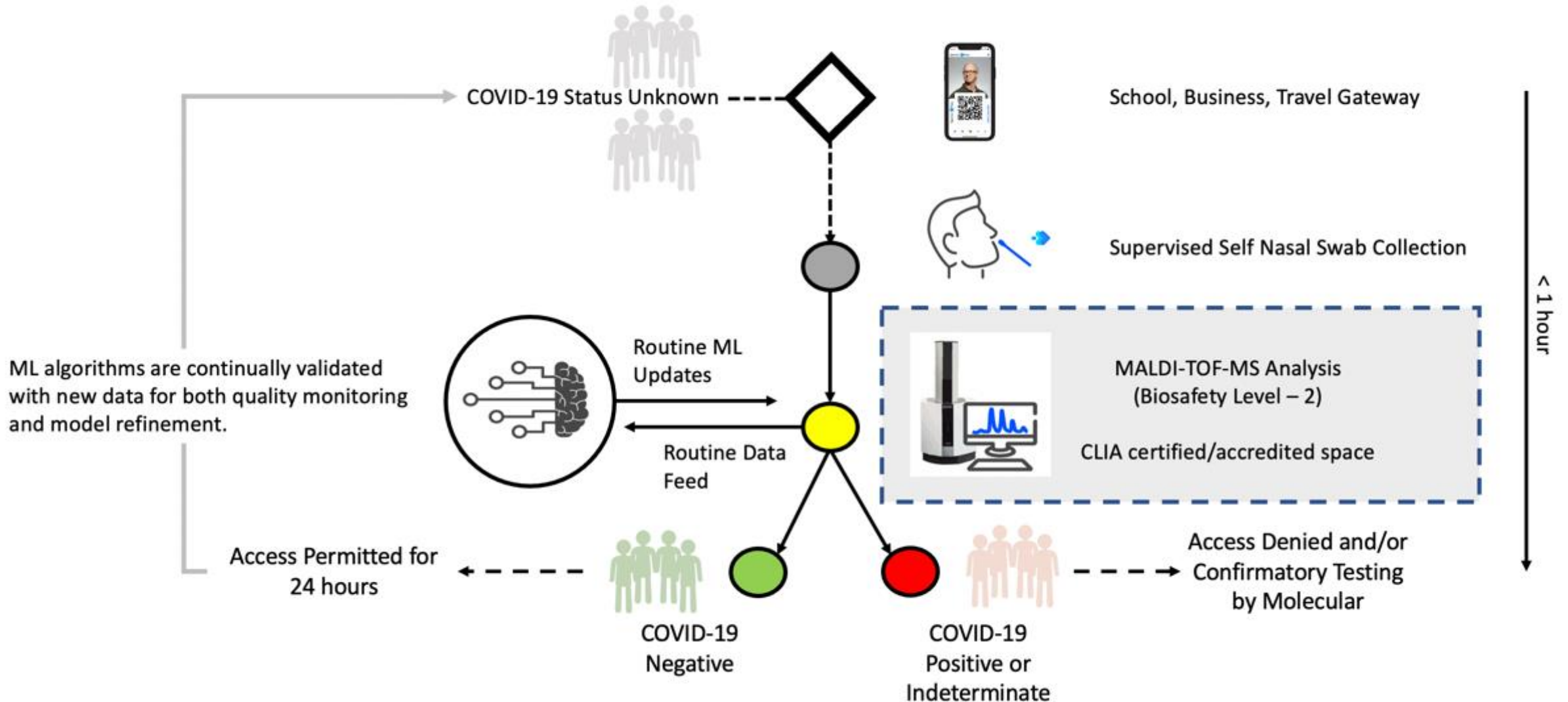
Proving to Ourselves this is Real – Further Generalization and Validation!

- MILO outperformed the competing autoML platform.
- Logistic regression produced the optimal platform:
 - AUC: 0.989
 - Accuracy: 92.8%
 - Sensitivity: 100%
 - Specificity: 90.3%

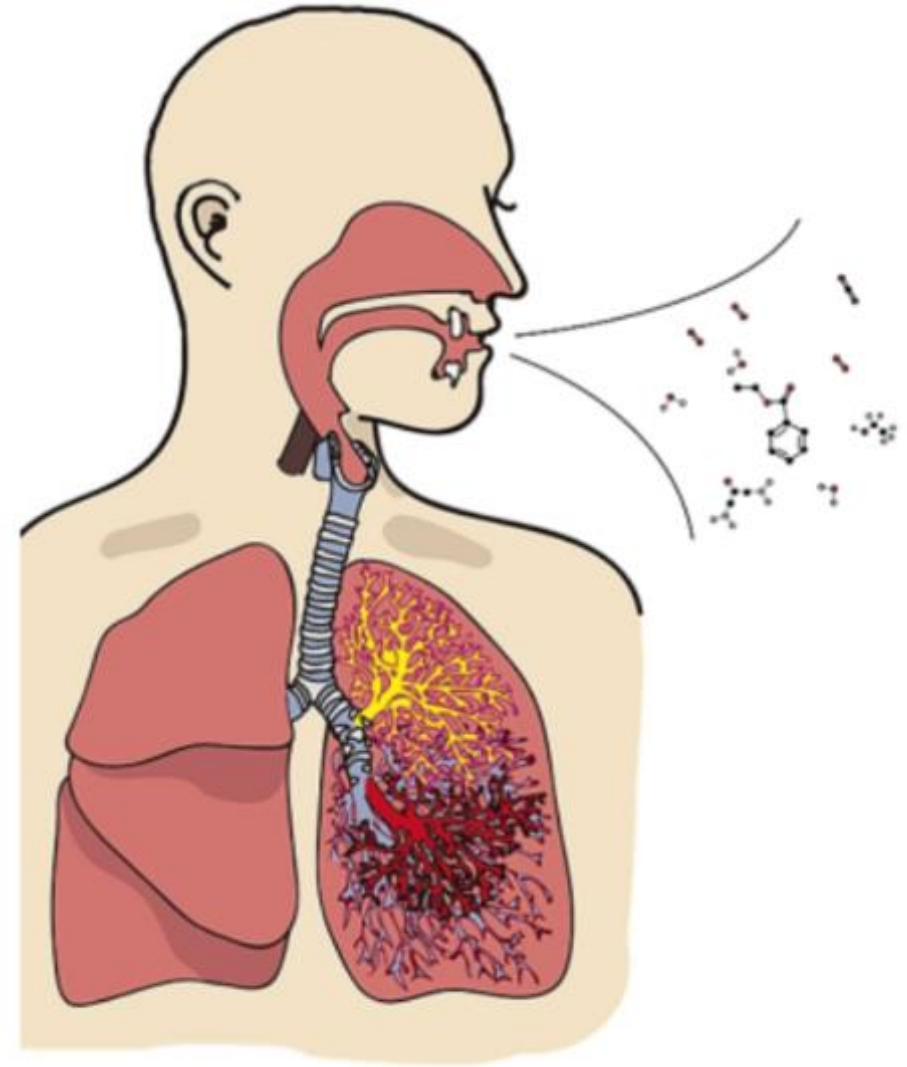
A. MILO <u>AutoML</u> generated Models						
Method	Accuracy % (95% CI)	AUROC (95% CI)	Positive Percent Agreement (PPA) % (95% CI)	Negative Percent Agreement (NPA) % (95% CI)	F1 Score	% features Selected
LBFGS-Logistic Regression	92.8 (88.2-96.0)	98.9 (81.9-100)	100 (92.9-100)	90.3 (84.2-94.6)	91.3	All*
k-Nearest Neighbor	92.3 (87.6-95.6)	96.9 (60.1-100)	100 (92.9-100)	89.6 (83.4-94.1)	90.7	25%#
Naïve Bayes	91.7 (86.9-95.2)	99.2 (84.8-100)	100 (92.9-100)	88.9 (82.6-93.5)	90.2	All*
Random Forest	95.4 (91.4-97.9)	98.1 (83.3-100)	92.0 (80.8-97.7)	96.5 (92.1-98.9)	93.9	All*
Support Vector Machine	93.3 (88.8-96.4)	98.6 (86.8-100)	100 (92.9-100)	91.0 (85.1-95.1)	91.9	75%##
Neural Network-Multi Layer Perceptron	94.9 (90.7-97.5)	99.6 (84.9-100)	100 (92.9-100)	93.1 (87.6-96.6)	92.5	All*
Gradient Boosting Machine (XGBoost)	93.8 (89.4-96.8)	98.3 (82.0-100)	94.0 (83.5-98.7)	93.8 (88.5-97.1)	92.2	All*

B. Microsoft <u>AutoML</u> generated Models						
Method	Accuracy % (95% CI)	AUROC **	Positive Percent Agreement (PPA) % (95% CI)	Negative Percent Agreement (NPA) % (95% CI)	F1 Score	% features Selected
Fast Tree	87.1 (81.6-91.5)	98.0	98.0 (89.4-99.9)	83.3 (76.2-89.0)	79.7	All
Fast Forest	86.6 (80.9-91.1)	96.9	92.0 (80.8-97.8)	84.7 (77.8-90.2)	78.0	All
Gradient Boosting Machine (light)	86.1 (80.4-90.6)	98.3	98.0 (89.4-99.9)	81.9 (74.7-87.9)	78.4	All
Support Vector Machine	95.4 (91.4-97.9)	99.5	98.0 (89.4-99.9)	94.4 (89.4-97.6)	91.6	All
SDCA-Logistic Regression	91.8 (86.9-95.2)	99.4	100 (92.9-100)	88.9 (82.6-93.5)	86.2	All
LBFGS-Logistic Regression	90.7 (85.7-94.4)	99.3	100 (92.9-100)	87.5 (80.9-92.4)	84.8	All
SGD-Calibrated	91.2 (86.3-94.8)	99.1	98.0 (89.4-99.9)	88.9 (82.6-93.5)	85.2	All
Symbolic SGD-Logistic Regression	85.6 (79.8-90.2)	97.1	92.0 (80.8-97.8)	83.3 (76.2-89.0)	76.7	All
Averaged Perceptron	89.2 (83.9-93.2)	98.7	98.0 (89.4-99.9)	86.1 (79.4-91.3)	82.4	All

Conceptual Workflow of a Machine Learning Enhanced MALDI-TOF-MS-Based Method for Detecting COVID-19:

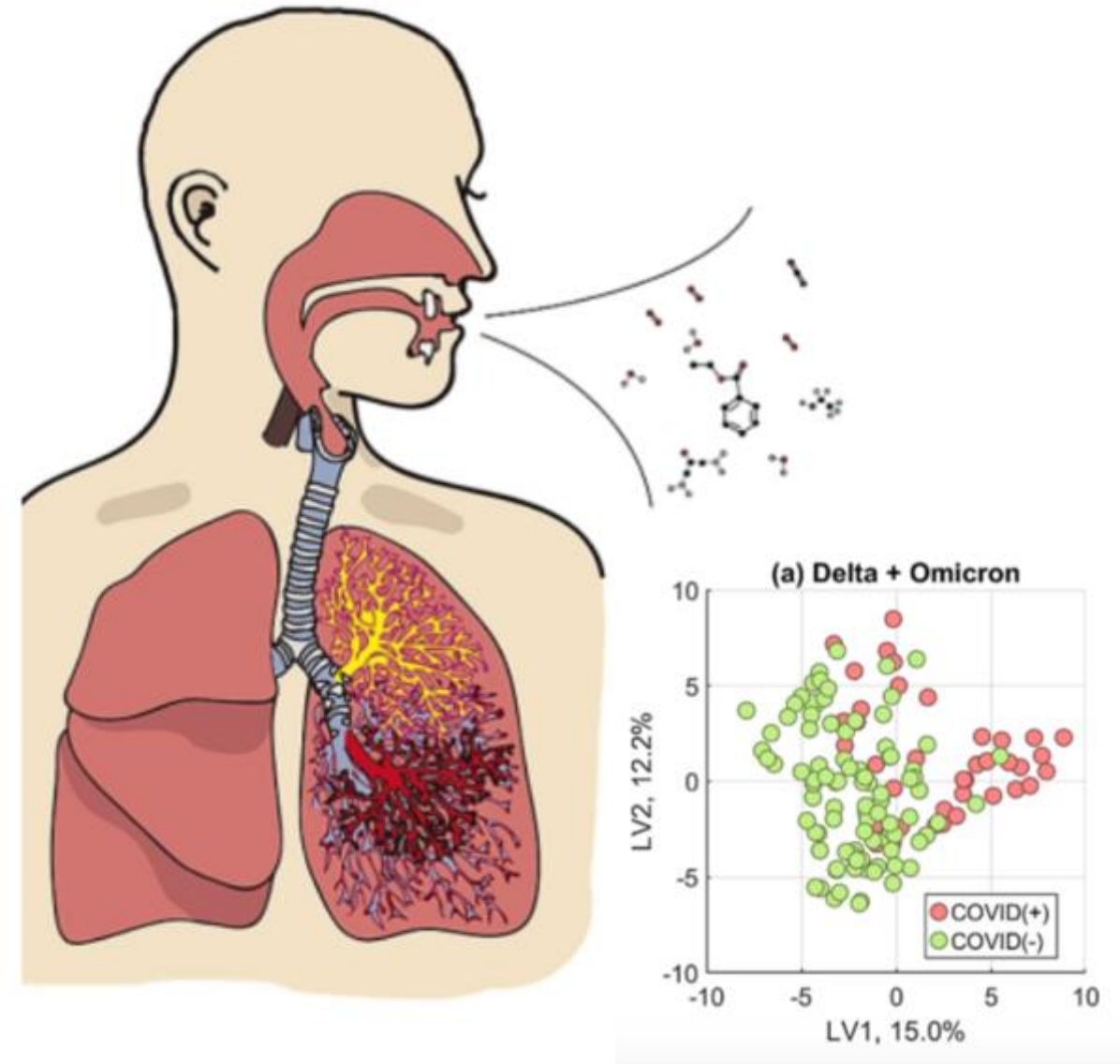


Can we use this ML-enhanced mass spectrometry for other sample types?



Breath Samples as a New Specimen Paradigm

- Volatile organic chemicals (VOC) exhaled by patients could be collected and analyzed to predict various diseases.
- When aided by machine learning, VOCs detected by mass spectrometry could be very specific and differentiate between infectious diseases.



Breath Samples as a New Specimen Paradigm

- Volatile organic chemicals (VOC) exhaled by patients could be collected and analyzed to predict various diseases.
- When aided by machine learning, VOCs detected by mass spectrometry could be very specific and differentiate between infectious diseases.
- Our team at UC Davis Health was able to show VOC (63 compounds in total) detection aided by machine learning could even differentiate between SARS-CoV-2 variants (accuracy 82-84%, specificity of 79%).

communications medicine

ARTICLE



<https://doi.org/10.1038/s43856-022-00221-5>

OPEN

Predominant SARS-CoV-2 variant impacts accuracy when screening for infection using exhaled breath vapor

Mitchell M. McCartney^{1,2,3,7}, Eva Borrás^{1,2,7}, Dante E. Rojas^{1,2}, Tristan L. Hicks^{1,2}, Katherine L. Hamera^{1,2}, Nam K. Tran⁴, Tina Tham⁵, Maya M. Juarez⁵, Enrique Lopez⁶, Nicholas J. Kenyon^{2,3,5} & Cristina E. Davis^{1,2,3}✉

Abstract

Background New technologies with novel and ambitious approaches are being developed to diagnose or screen for SARS-CoV-2, including breath tests. The US FDA approved the first breath test for COVID-19 under emergency use authorization in April 2022. Most breath-based assays measure volatile metabolites exhaled by persons to identify a host response to infection. We hypothesized that the breathprint of COVID-19 fluctuated after Omicron became the primary variant of transmission over the Delta variant.

Methods We collected breath samples from 142 persons with and without a confirmed COVID-19 infection during the Delta and Omicron waves. Breath samples were analyzed by gas chromatography-mass spectrometry.

Results Here we show that based on 63 exhaled compounds, a general COVID-19 model had an accuracy of 0.73 ± 0.06 , which improved to 0.82 ± 0.12 when modeling only the Delta wave, and 0.84 ± 0.06 for the Omicron wave. The specificity improved for the Delta and Omicron models (0.79 ± 0.21 and 0.74 ± 0.12 , respectively) relative to the general model (0.61 ± 0.13).

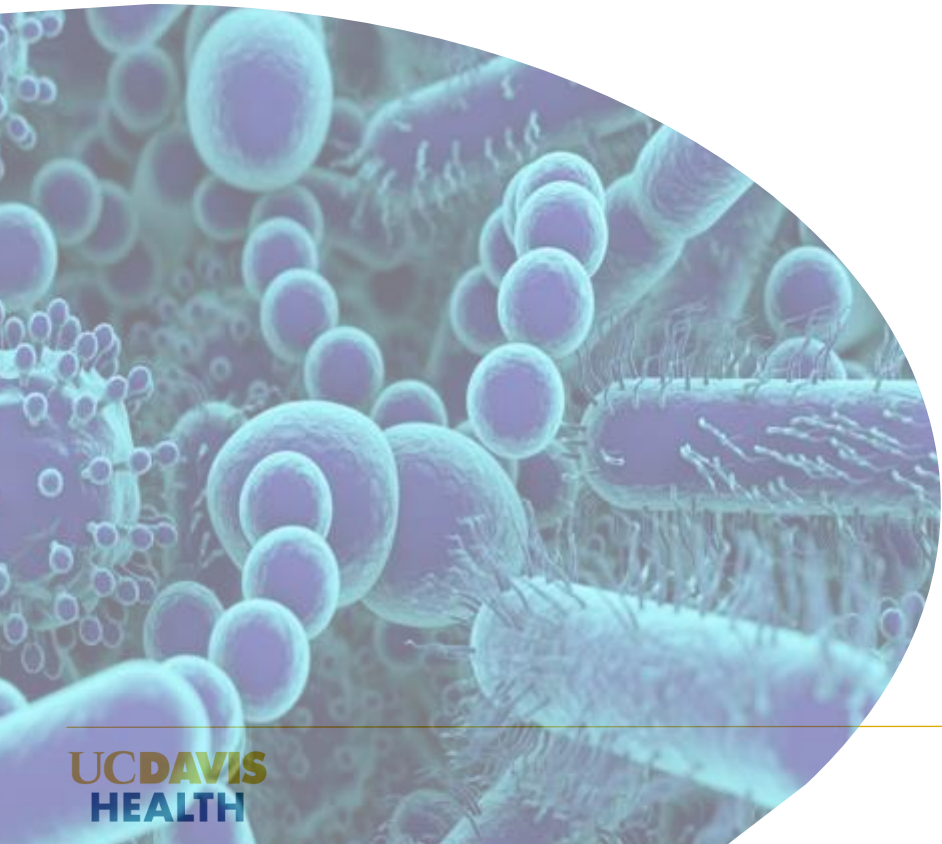
Conclusions We report that the volatile signature of COVID-19 in breath differs between the Delta-predominant and Omicron-predominant variant waves, and accuracies improve when samples from these waves are modeled separately rather than as one universal approach. Our findings have important implications for groups developing breath-based assays for COVID-19 and other respiratory pathogens, as the host response to infection may significantly differ depending on variants or subtypes.

Plain language summary

In recent decades, scientists have found we exhale thousands of compounds that reveal much about our health, including whether we are sick with COVID-19. Our team asked whether the breath profile of someone infected with the Delta variant of COVID-19 would match the breath profile caused by the Omicron variant—a version of the virus that is more transmissible. We analyzed breath samples from 142 people, some sick with either the Delta or Omicron variant of COVID-19, and others who were negative for COVID-19. Our results indicate that the Delta variant altered the contents of our breath in a different way than the Omicron variant, and breath-based tests improved when optimized to detect only one of the variants. These findings might impact the design of future breath-based tests for COVID-19.

Sepsis: The Clinical Problem

- Sepsis is defined as life-threatening organ dysfunction caused by a dysregulated host response to infection
- Over 750,000 patients in the United States experience sepsis each year.
- Mortality ranges from 28-50% and can be as high as 90% in cases of septic shock.



Sepsis



Breast cancer, prostate cancer, and AIDS



Multicenter Burn Sepsis Data



Overview:

- Burn sepsis is hard to detect and current criteria exhibit poor sensitivity and specificity.
- Dataset derived from an American Burn Association / Department of Defense funded sepsis study (n = 218 patients).

ClinicalTrials.gov Identifier: NCT01140269

Tran NK, et al. Sci Rep 2020;10:12354

Multicenter Burn Sepsis Data

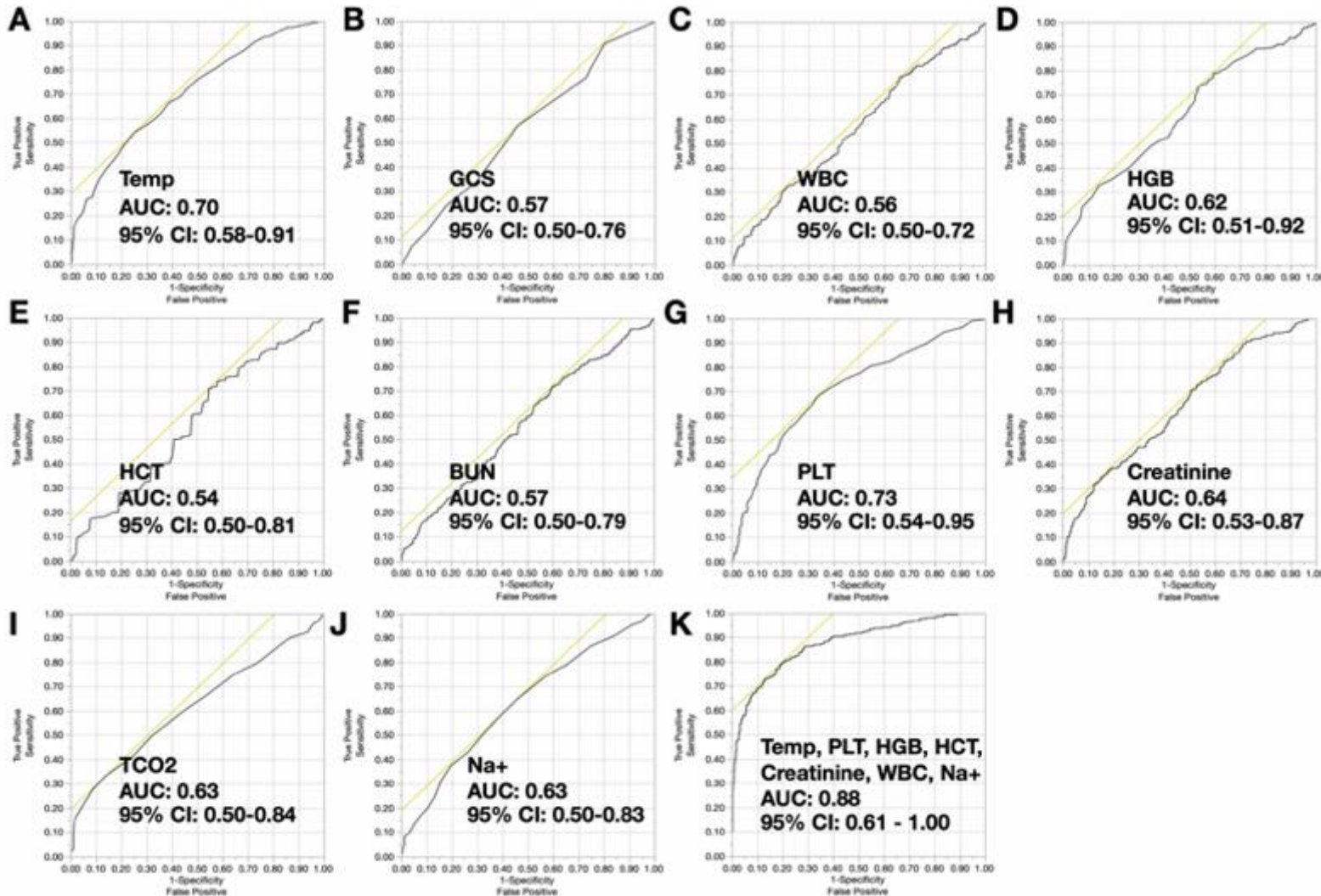
Test Category	Test Name
Chemistry	BMP, CMP
Coagulation	INR, aPTT
Hematology	CBC w/ auto diff
Microbiology	Bacterial and fungal culture, plus MALDI-TOF-MS
Acid-Base	Blood gases
Molecular	PCR pathogen identification

Overview:

- Burn sepsis is hard to detect and current criteria exhibit poor sensitivity and specificity.
- Dataset derived from an American Burn Association / Department of Defense funded sepsis study (n = 218 patients).
- Daily vitals and routine lab results were collected. Plus PCR infectious disease data.

*Plus vital signs paired to laboratory results

Burn Sepsis Prediction with Traditional Statistics



American Burn Association (ABA) Sepsis Criteria:

Temp: >39 C or <36.5 C

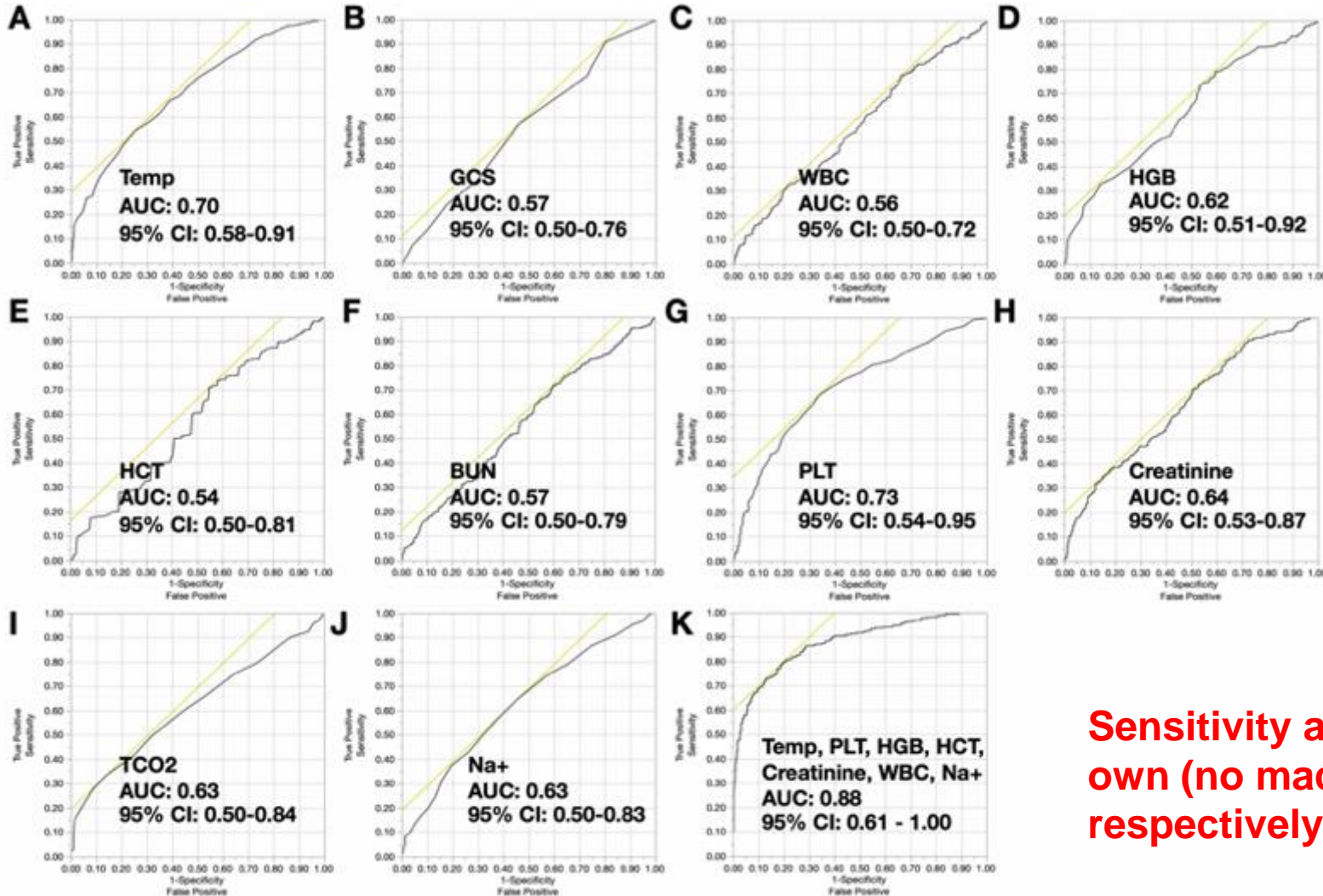
Heart Rate: >110 beats/min

Respiratory: >25 breaths/min

WBC: Not used

Platelet Count: <100,000 cells/uL

Burn Sepsis Prediction with Traditional Statistics



American Burn Association (ABA) Sepsis Criteria:

Temp: >39 C or <36.5 C

Heart Rate: >110 beats/min

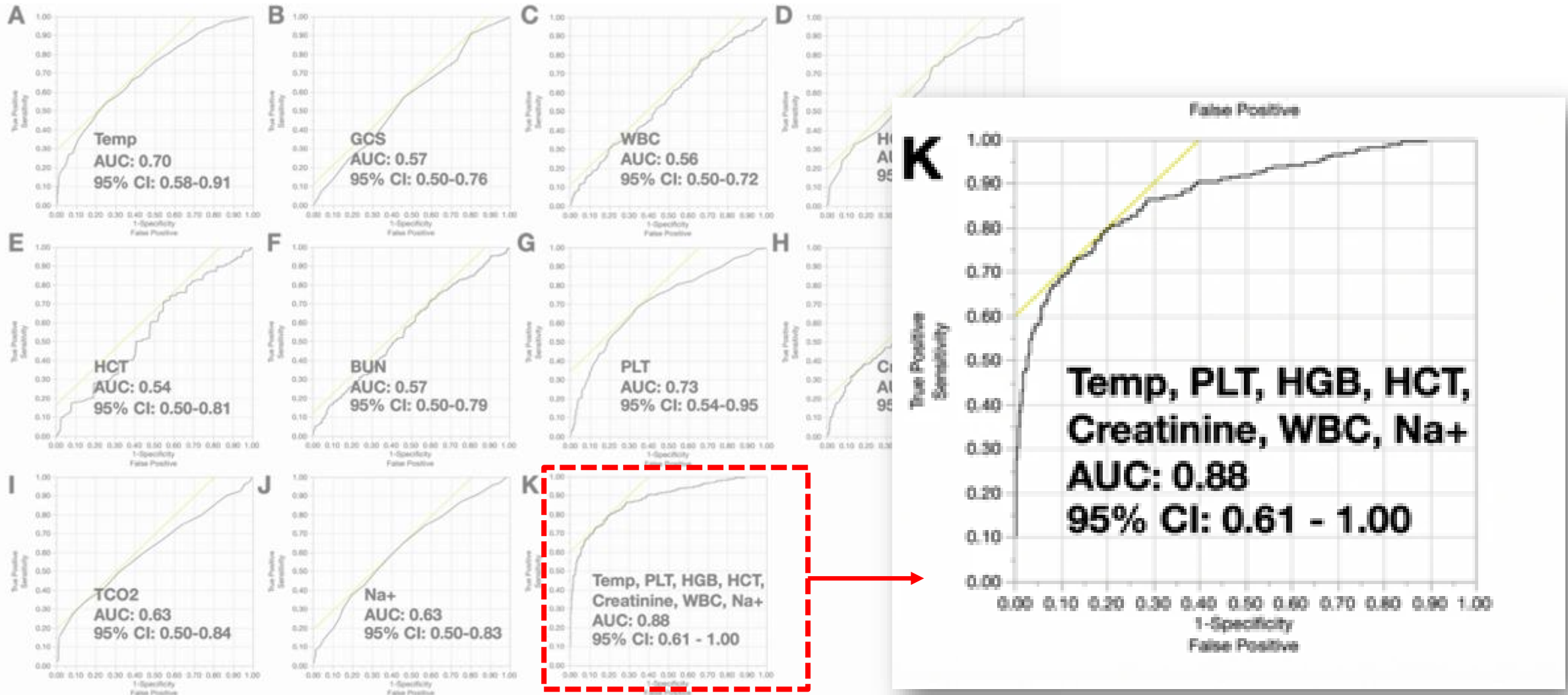
Respiratory: >25 breaths/min

WBC: Not used

Platelet Count: <100,000 cells/uL

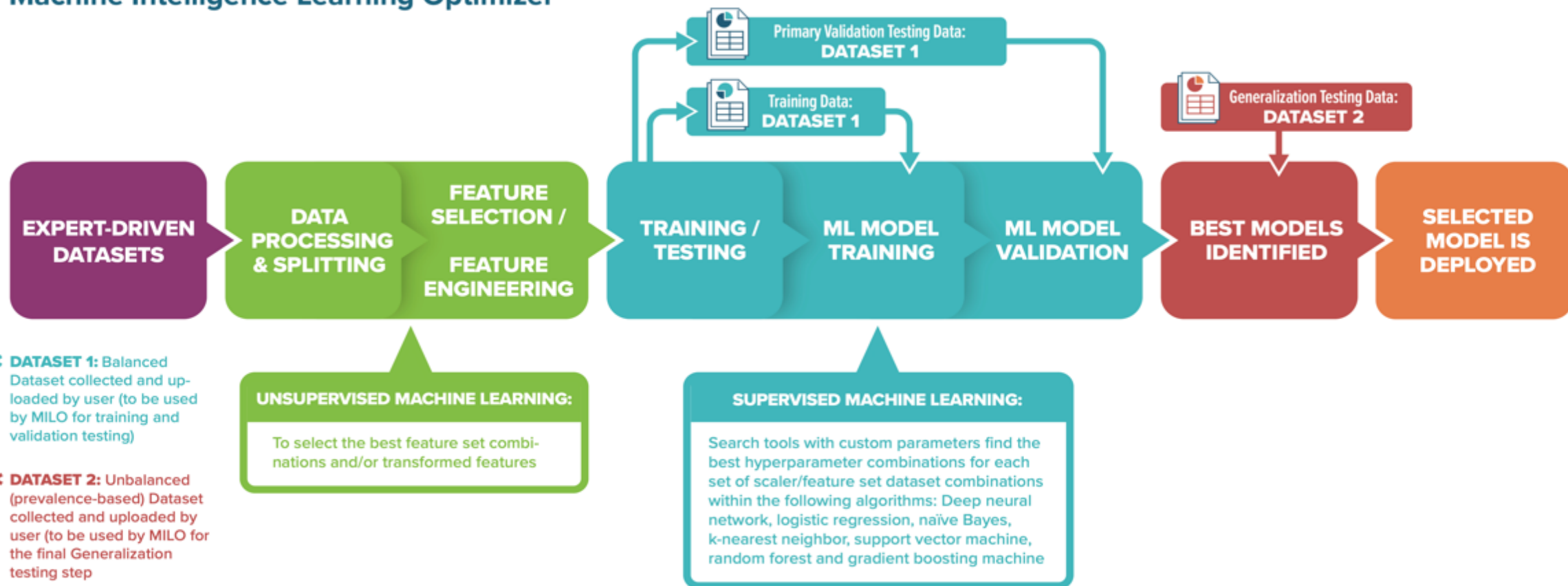
Sensitivity and Specificity of ABA Criteria on its own (no machine learning) is 75.0% and 65.7% respectively!

Burn Sepsis Prediction with Traditional Statistics



MILO AUTO-ML PLATFORM

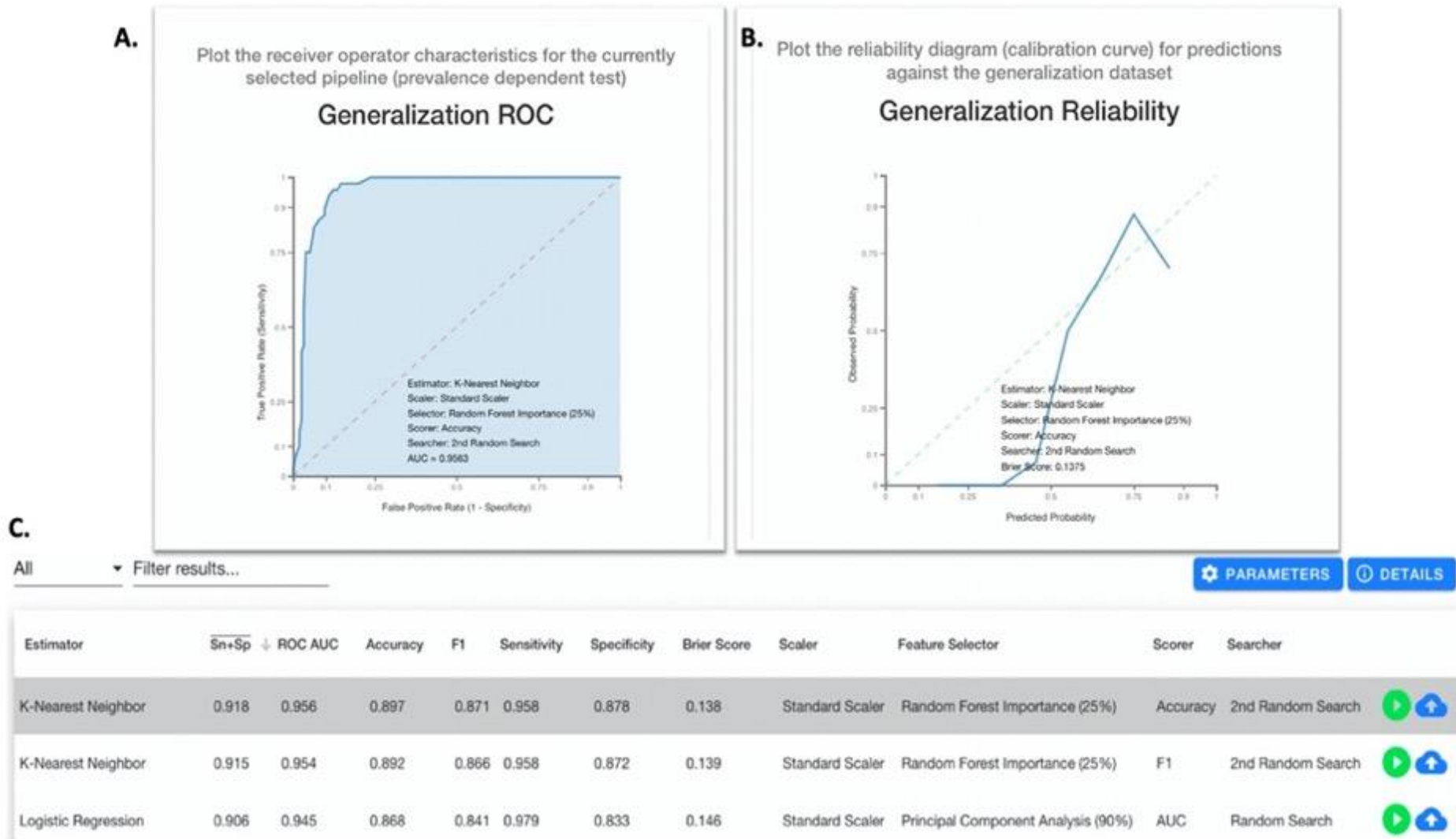
Machine Intelligence Learning Optimizer



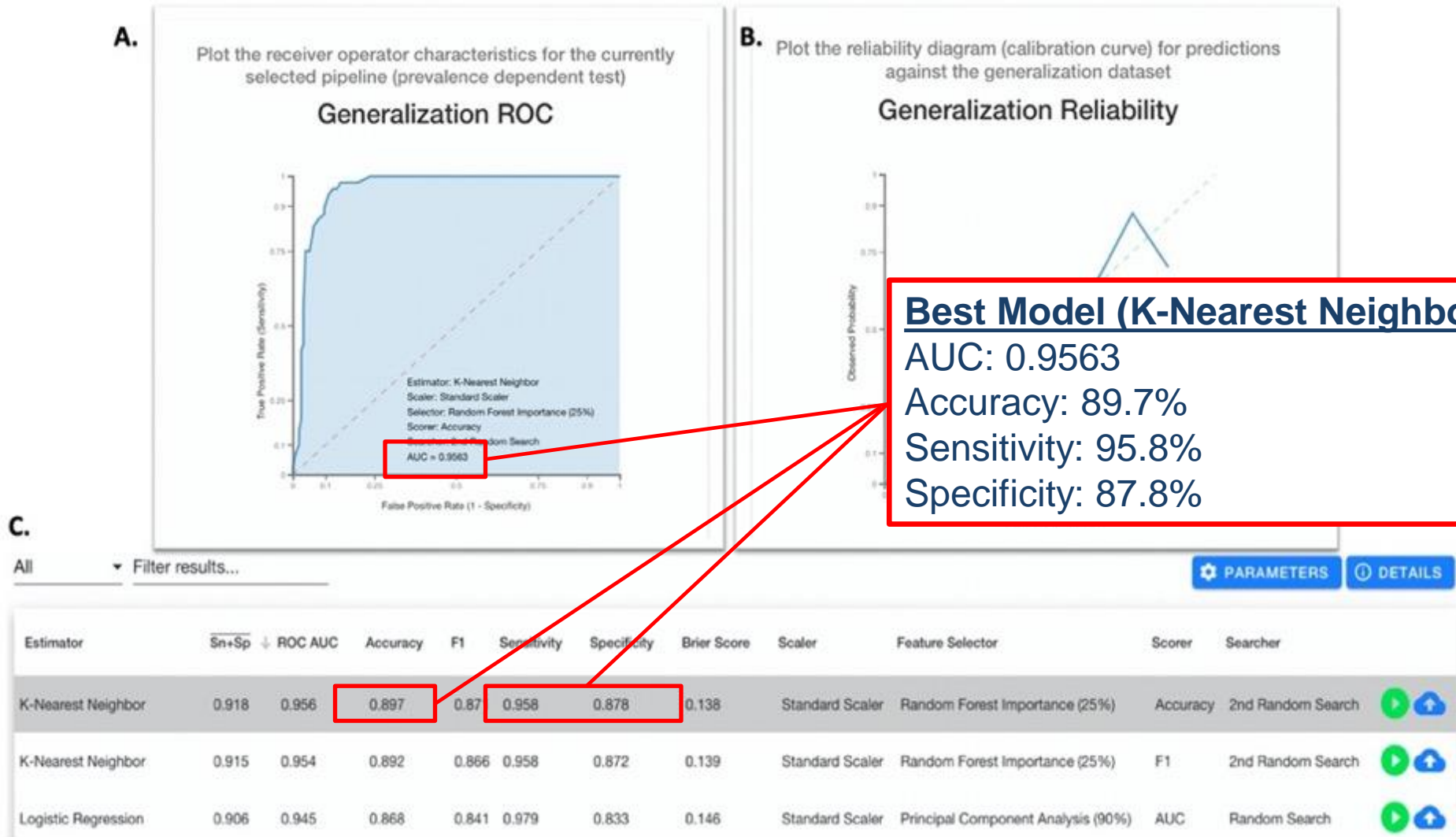
NOTES

- MILO removes missing values in Datasets automatically
- Dataset's number and percentages within each category is automatically assessed
- Data is scaled per pipeline using of variety of scaling methods
- MILO automatically builds and deploys the user-selected model in a WebApp (this live model can now be used to test new data and make predictions)
- The PMML or joblib file can also be easily exported as per user's needs

Sepsis Prediction with MLO Algorithm(s)



Sepsis Prediction with MILO Algorithm(s)



Best Model (K-Nearest Neighbor)

AUC: 0.9563

Accuracy: 89.7%

Sensitivity: 95.8%

Specificity: 87.8%

Example Comparing AI/ML Performance versus Traditional Sepsis Criteria for Burn Sepsis

	MILO AI/ML	ABA Criteria	SEPSIS-3
Features (Predictors)	Hgb, BUN, TCO2, temperature, heart rate	Temperature, heart rate, platelet count, respiratory rate	SOFA score parameters
Area Under the ROC	0.96	0.76	0.55
Sensitivity (%)	95.8	75.0	61.2
Specificity (%)	87.8	65.7	55.1

Abbreviations: ABA, American Burn Association; SOFA, sequential organ failure assessment score.

SOFA Score: Glasgow Coma Score, blood pressure, PaO2/FiO2 ratio, creatinine, platelets, total bilirubin, respiratory rate.

AI/ML Laboratory Opportunities

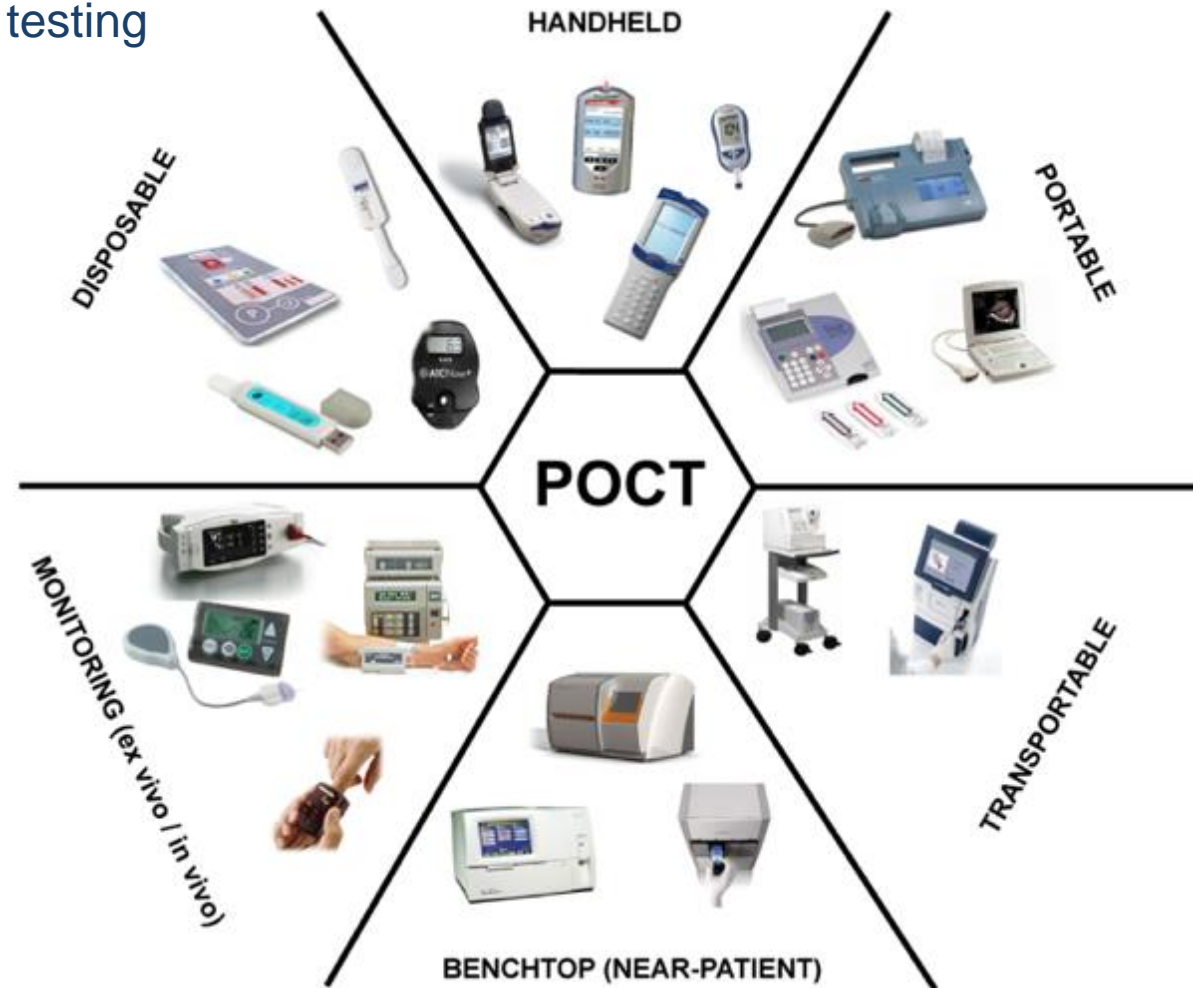
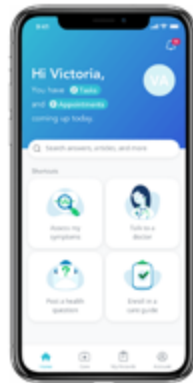
OPPORTUNITY	EXAMPLES
Error Detection / Prevention	Specimen integrity, mislabeled specimens, facial recognition
Image / Pattern recognition	Slide and fluid analysis
Multi-Analyte / Complex Data Analysis	Mass spectrometry, “big data” applications (e.g., EHR data, genomics, etc)
Automated medical decision-making	Point-of-care testing

How about AI/ML for Point-of-Care Testing?

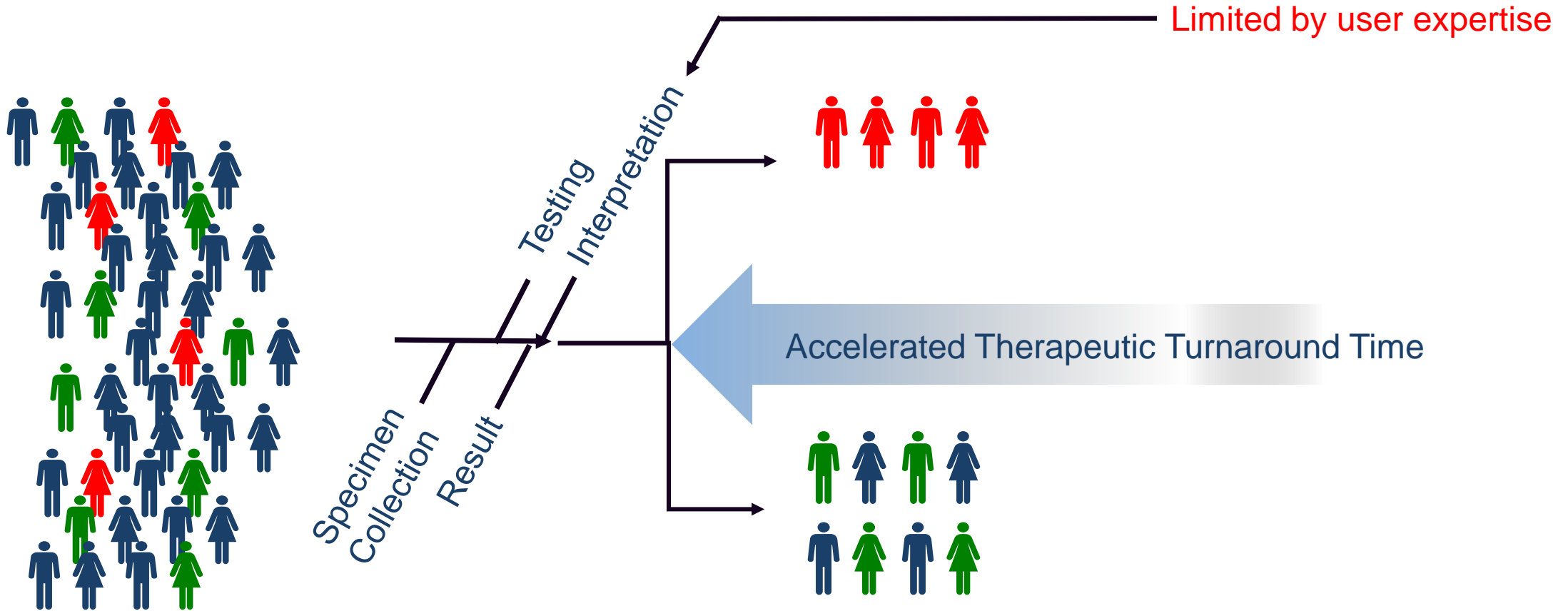
Point-of-Care Testing (POCT) is defined as medical testing at or near the site of patient care

Includes:

- Disposable
- Handheld
- Portable
- Transportable
- Benchtop
- Monitoring
- **Smart devices**



Could AI/ML be used to augment POCT result interpretation?



POINT-OF-CARE DIAGNOSTIC WORKFLOW

Burn-Related Acute Kidney Injury

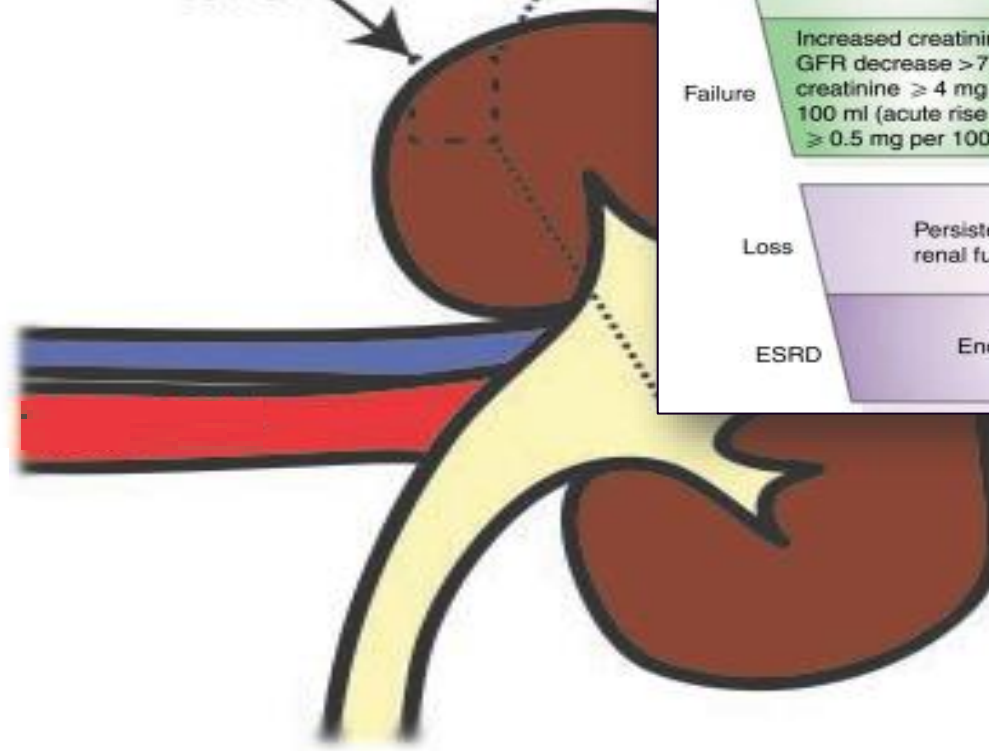
KDIGO AKI Criteria

	GFR criteria	Urine output criteria	
Risk	Increased creatinine $\times 1.5$ or GFR decrease $> 25\%$	UO $< 0.5 \text{ ml kg}^{-1} \text{ h}^{-1}$ $\times 6 \text{ h}$	High sensitivity
Injury	Increased creatinine $\times 2$ or GFR decrease $> 50\%$	UO $< 0.5 \text{ ml kg}^{-1} \text{ h}^{-1}$ $\times 12 \text{ h}$	
Failure	Increased creatinine $\times 3$ or GFR decrease $> 75\%$ or creatinine $\geq 4 \text{ mg per 100 ml}$ (acute rise of $\geq 0.5 \text{ mg per 100 ml dl}$)	UO $< 0.3 \text{ ml kg}^{-1} \text{ h}^{-1}$ $\times 24 \text{ h}$ or anuria $\times 12 \text{ h}$	Oliguria High specificity
Loss	Persistent ARF = complete loss of renal function $> 4 \text{ weeks}$		
ESRD	End-stage renal disease		

PRE-RENAL
Heart Failure
Hypotension
Sepsis



INTRINSIC
Acute Interstitial Nephritis
Hypoxic/Ischemic Injury
Nephrotoxic Injury



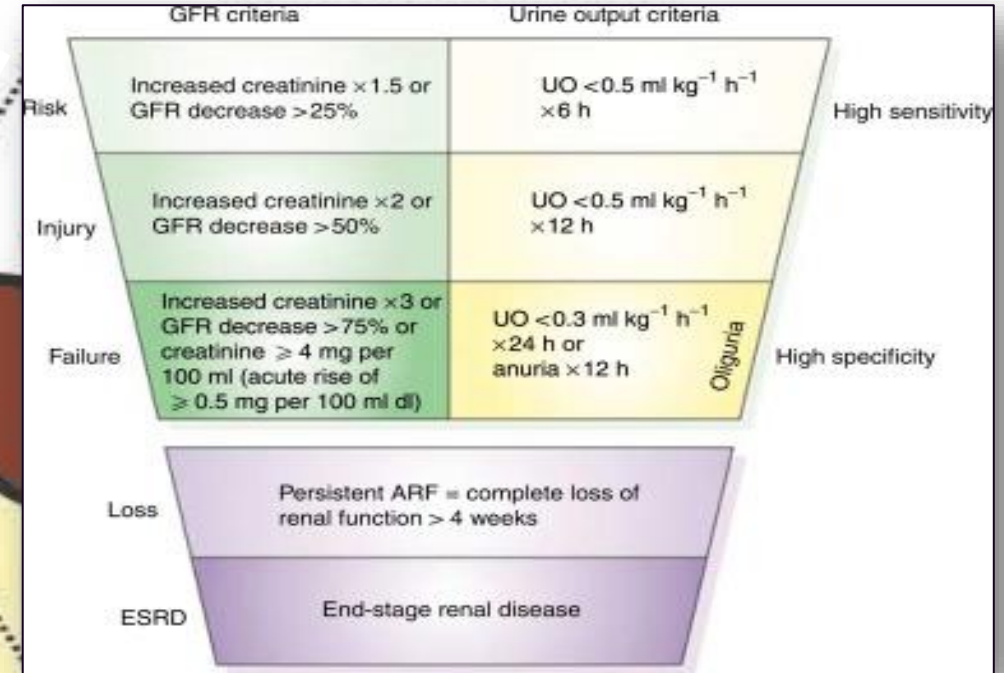
Up to 58% burn patients experience AKI. Causes:

- Sepsis
- Nephrotoxic medications
- “Burn Shock”

Biomarkers for Detecting Acute Kidney Injury

Biomarker	Pros	Cons
Creatinine	Easy to obtain Common test Familiarity	Half-life is 4 hours, requires – changes may not be discernable for up to 12 hours Up to 70% of nephrons can be damaged to see any real change.
Urine Output	Easy to obtain Common test Familiarity	AKI may already be occurring before urine output decreases
GFR	Easily estimated by creatinine values	GFR may not change too much early on due to complex autoregulation.

KDIGO AKI Criteria

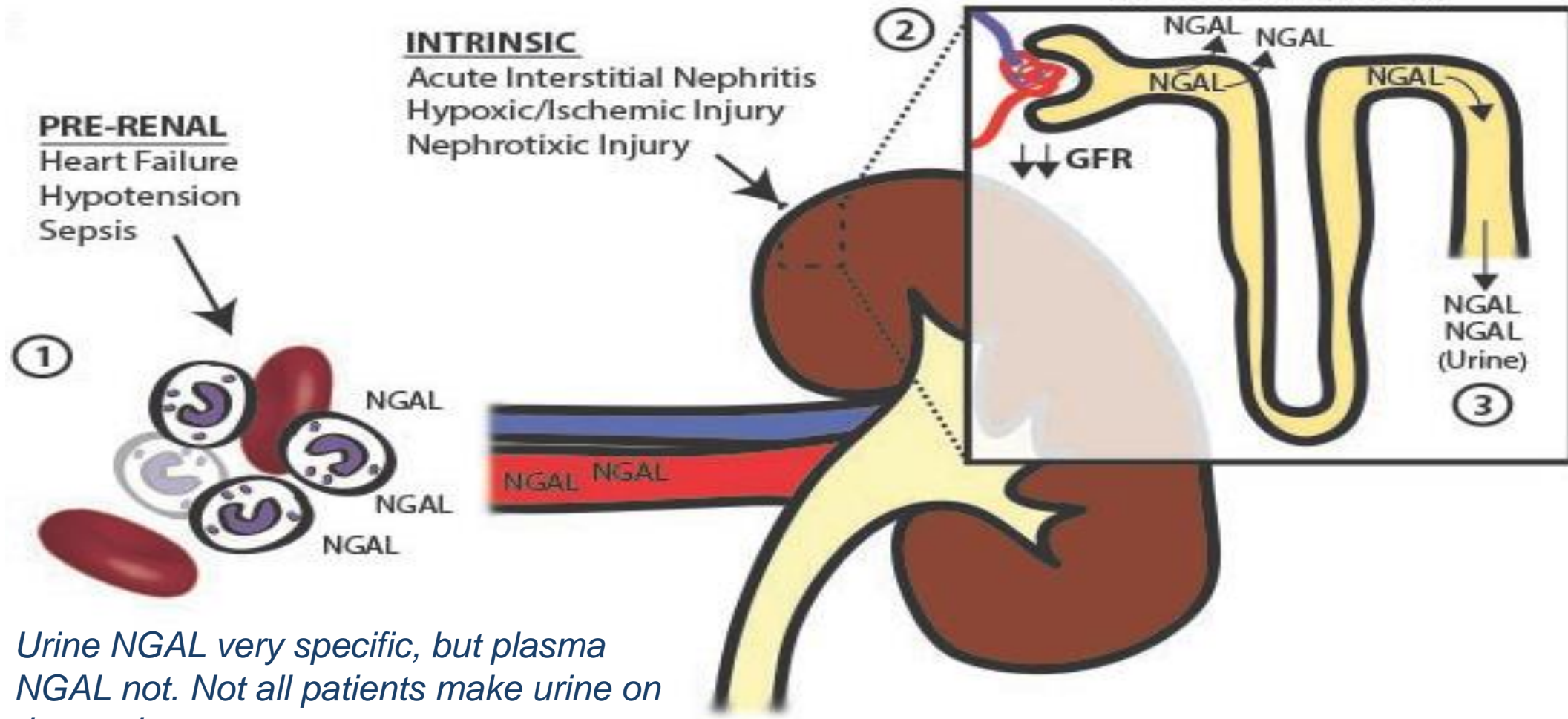


Up to 58% burn patients experience AKI. Causes:

- Sepsis
- Nephrotoxic medications
- “Burn Shock”

Biomarkers for Detecting Acute Kidney Injury

Neutrophil Gelatinase Associated Lipocalin

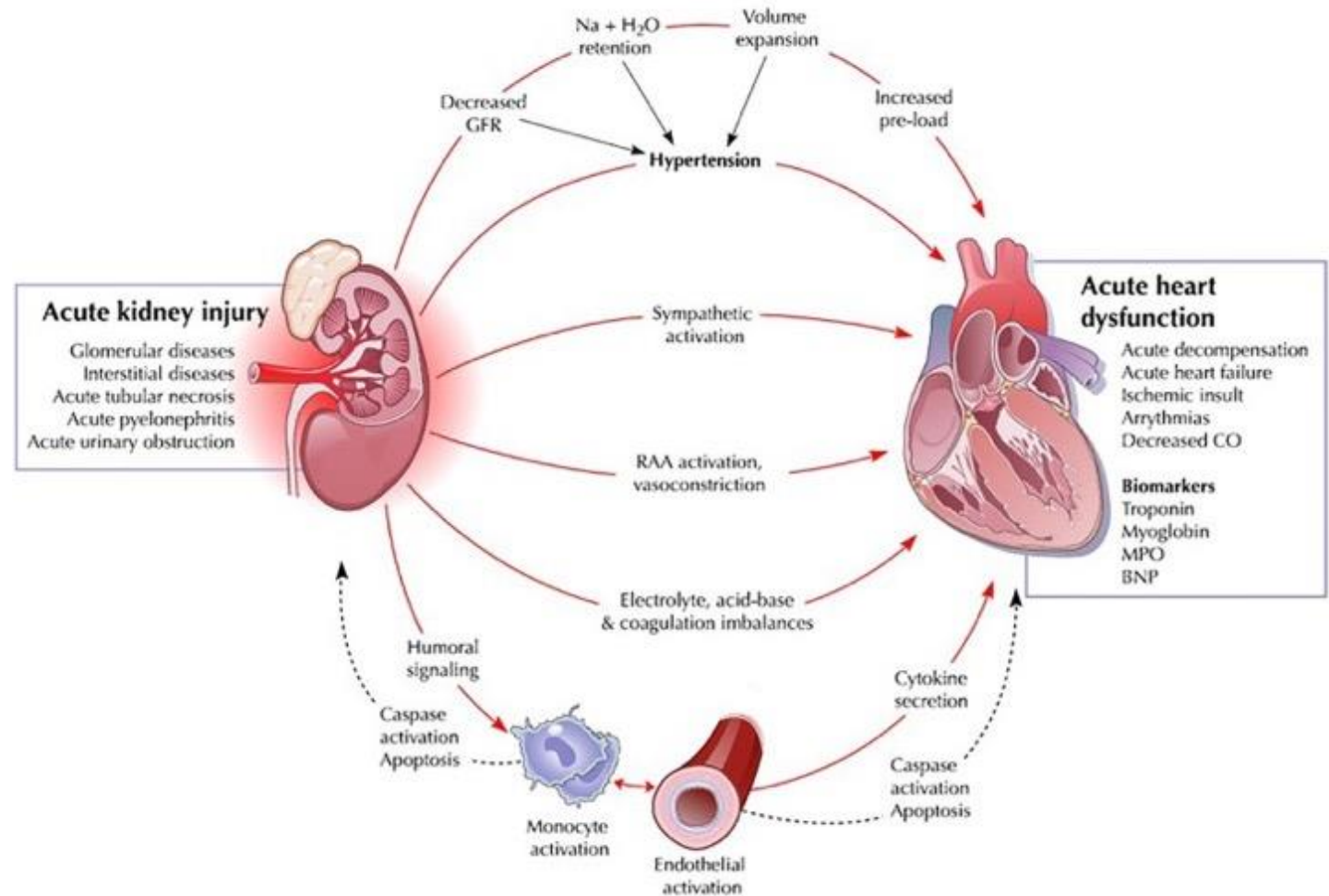


Urine NGAL very specific, but plasma NGAL not. Not all patients make urine on demand...

Biomarkers for Detecting Acute Kidney Injury

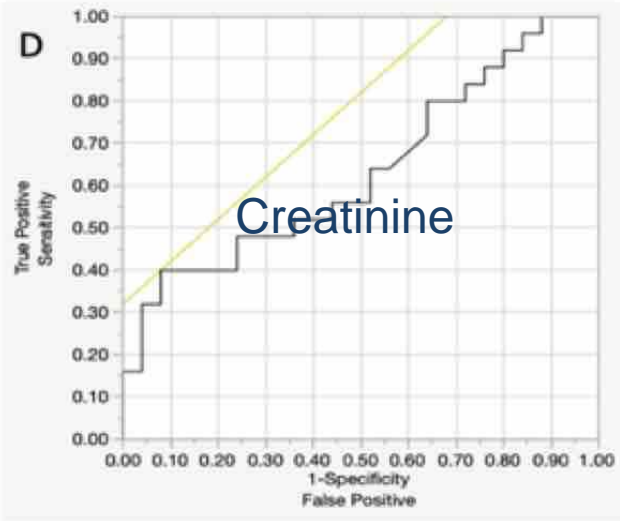
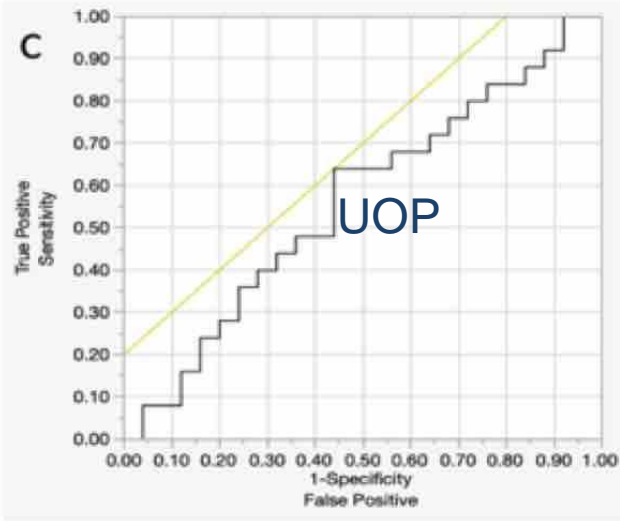
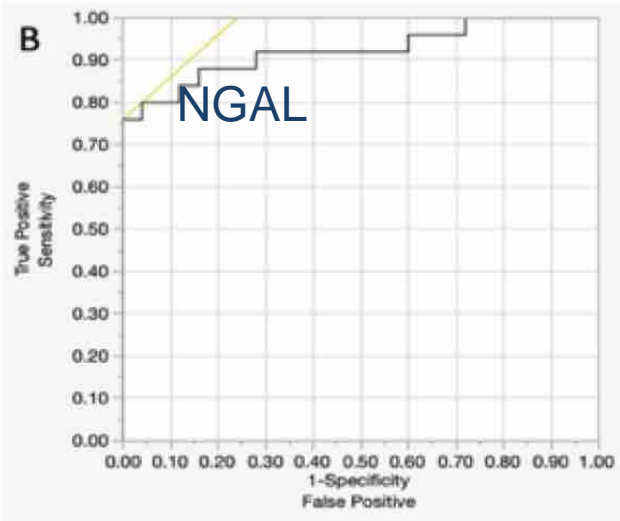
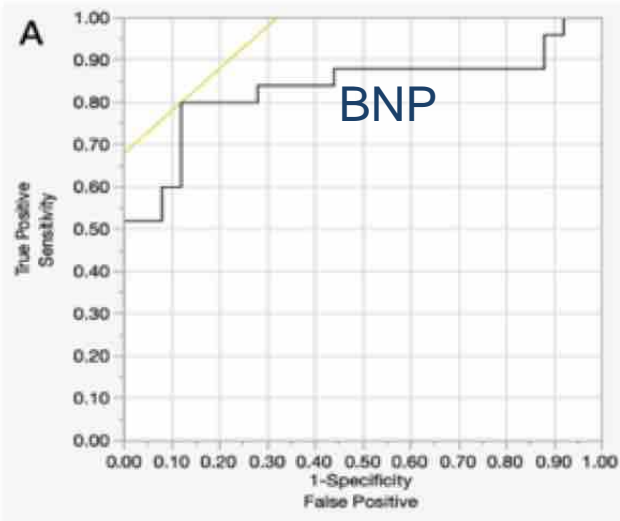
Natriuretic Peptides for Cardio-Renal Syndrome

- Acute kidney injury can result in acute heart dysfunction.
- Heart dysfunction results in the release of natriuretic peptides such as BNP and NT-proBNP.
- The use of **both plasma NGAL and BNP/NT-proBNP** improves prediction of AKI.
- Without BNP, plasma NGAL specificity was about 90.9% with an area under the ROC curve of 0.82.

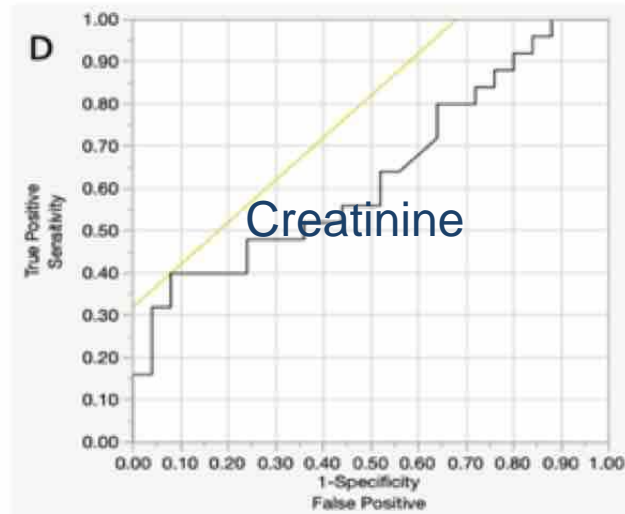
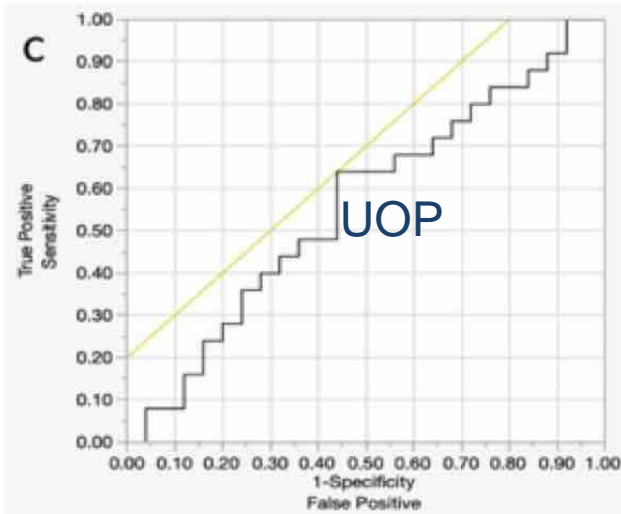
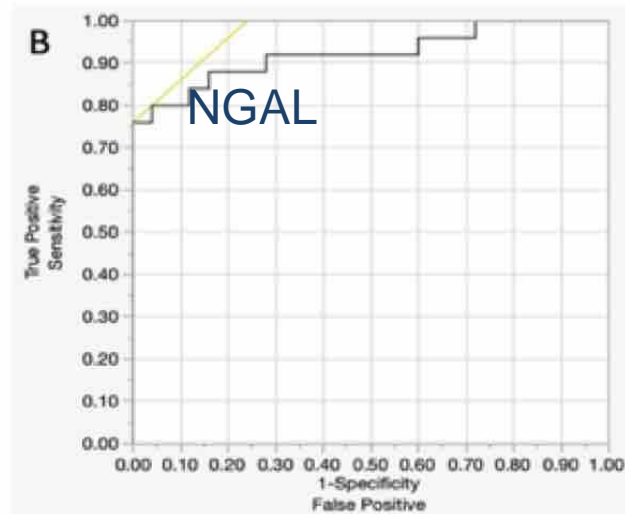
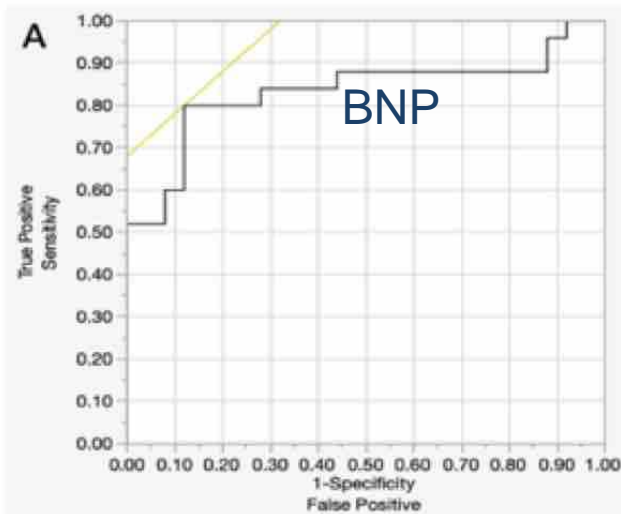


Clinical Performance of AKI Biomarkers

Receiver Operator Characteristic (ROC) Curves for AKI Biomarkers



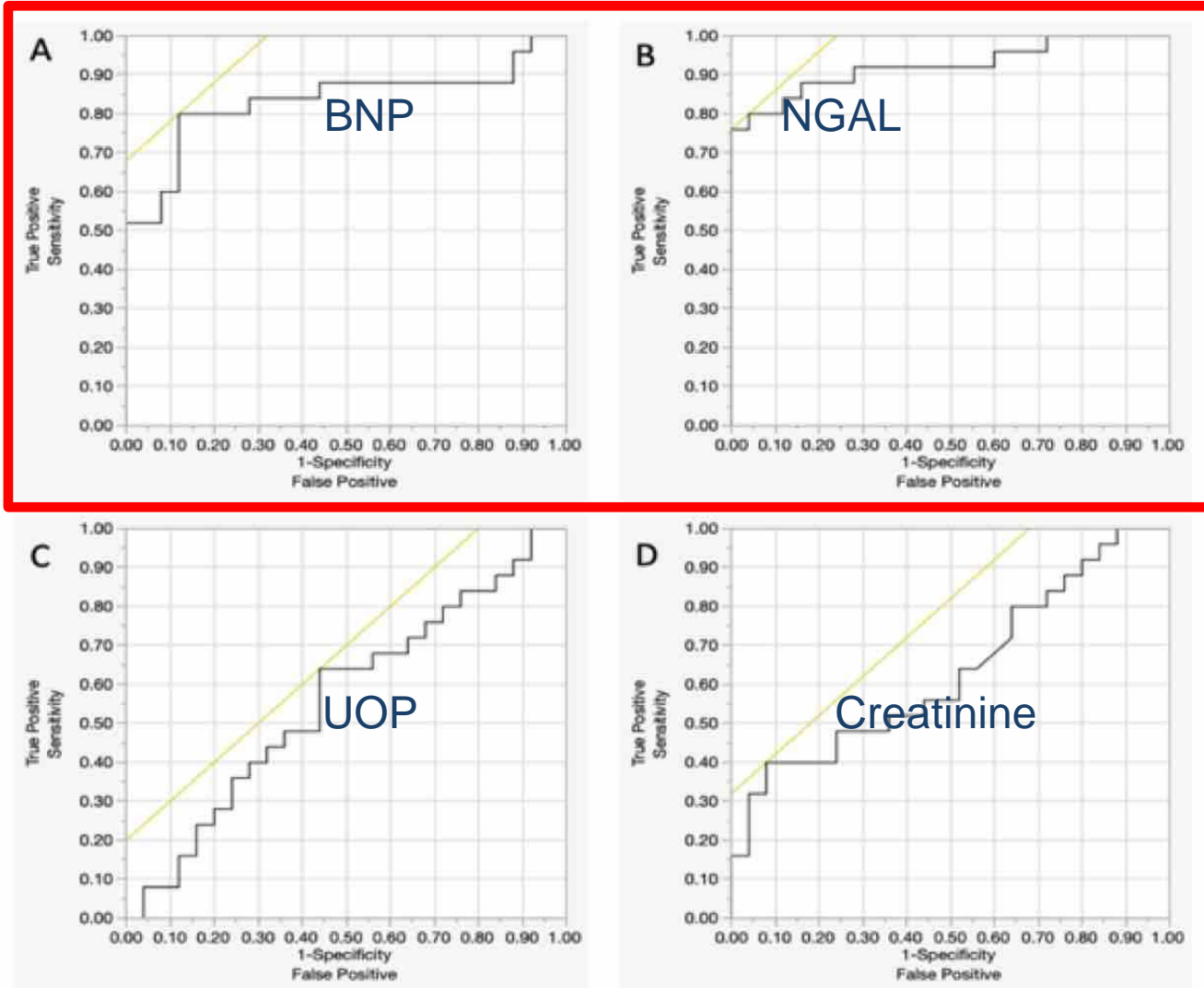
Clinical Performance of AKI Biomarkers



Receiver Operator Characteristic (ROC) Curves for AKI Biomarkers

- UOP and creatinine do not perform well at all, despite being the most common measurements to detect AKI.

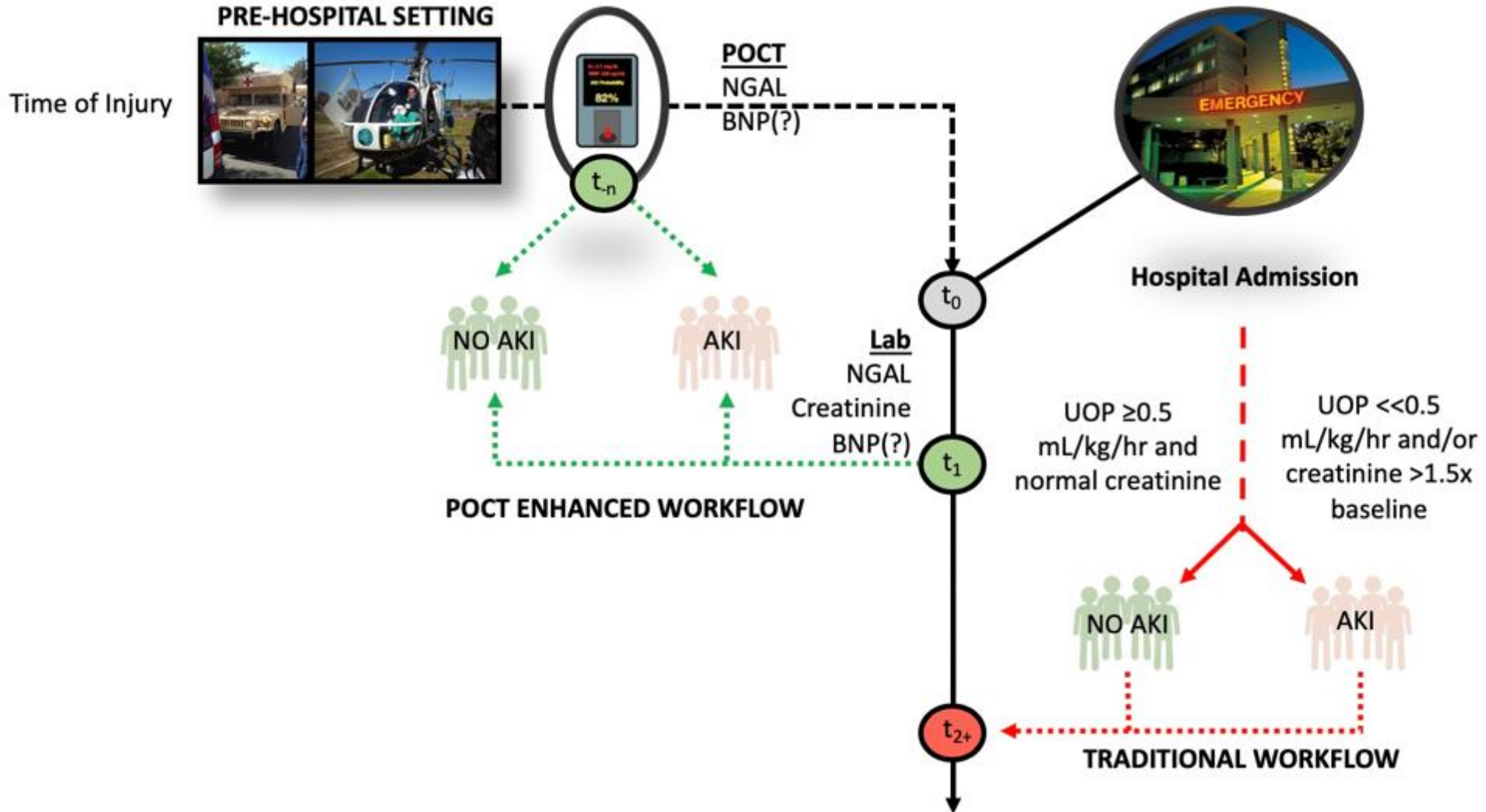
Clinical Performance of AKI Biomarkers



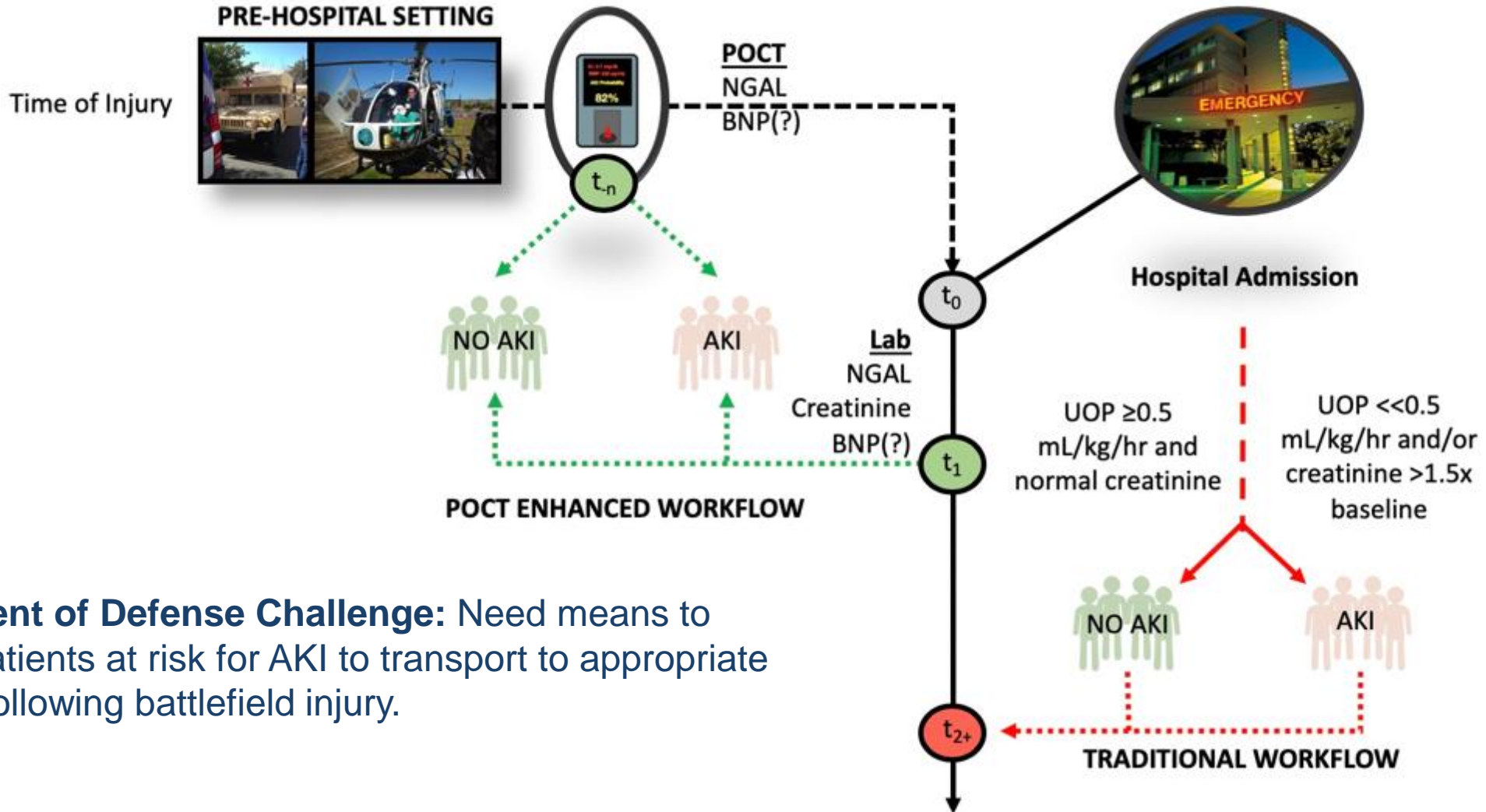
Receiver Operator Characteristic (ROC) Curves for AKI Biomarkers

- BNP and NGAL appear to perform better than UOP and creatinine.
- *So why not try to see if AI/ML can improve performance of these four biomarkers?*

AI/ML Enhanced Point-of-Care AKI Detection



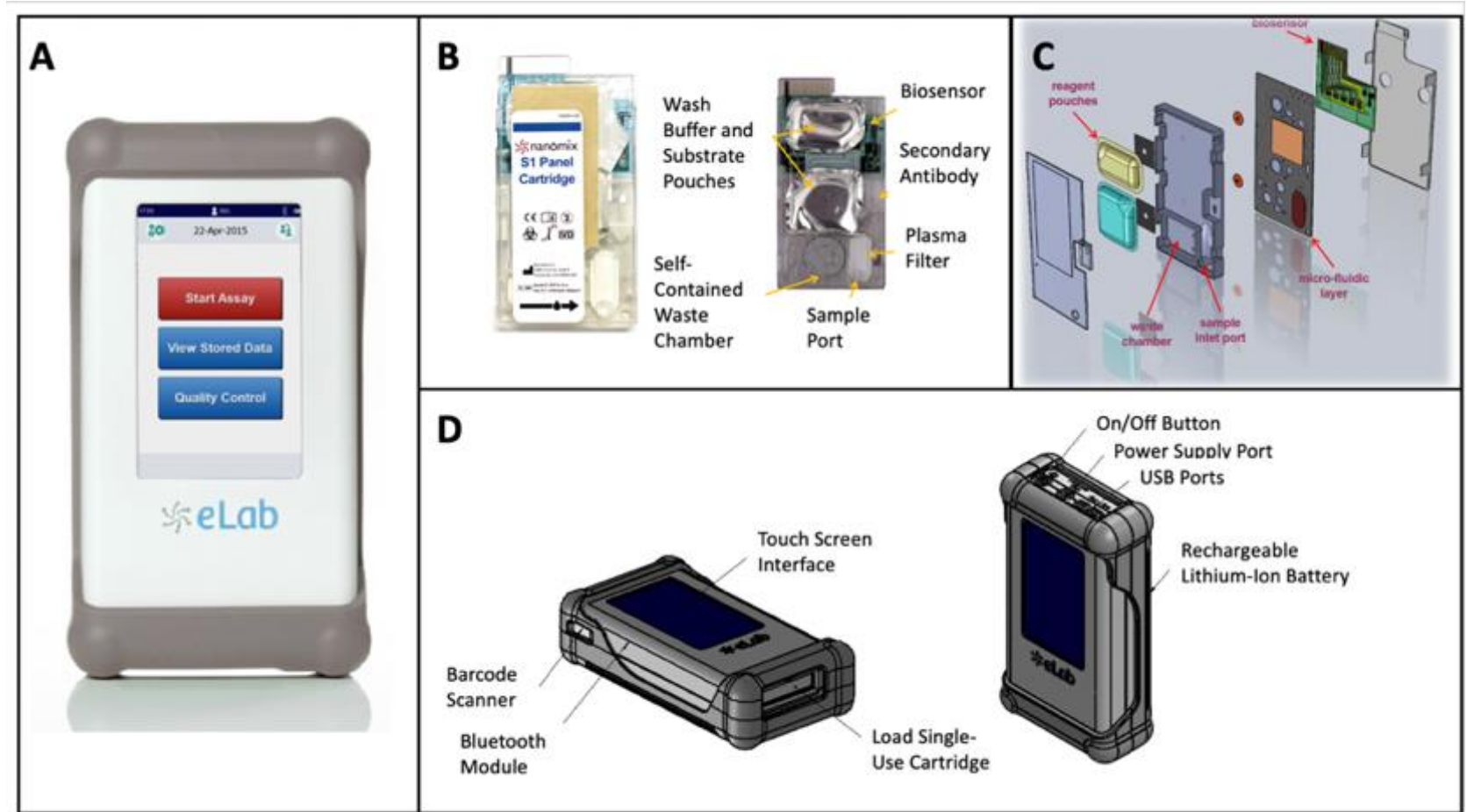
AI/ML Enhanced Point-of-Care AKI Detection



Department of Defense Challenge: Need means to identify patients at risk for AKI to transport to appropriate facilities following battlefield injury.

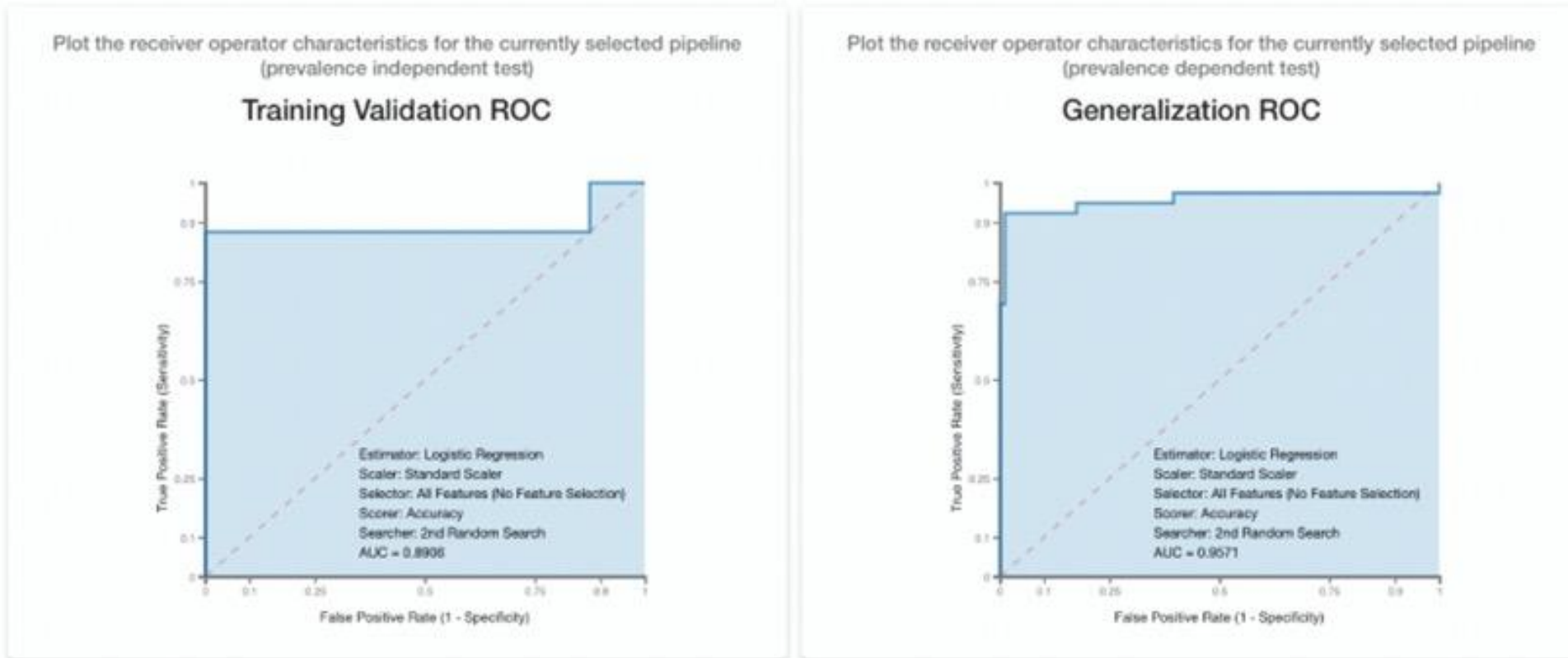
AI/ML Enhanced Point-of-Care AKI Detection

- Funded by DoD to develop POC assay for NGAL.
- Multicenter study with UC Davis and University of Cincinnati.
- Studied both burned and non-burned trauma patients.
- Machine learning employed to analyze data to enable battlefield decision making.



AI/ML Enhanced Point-of-Care AKI Detection

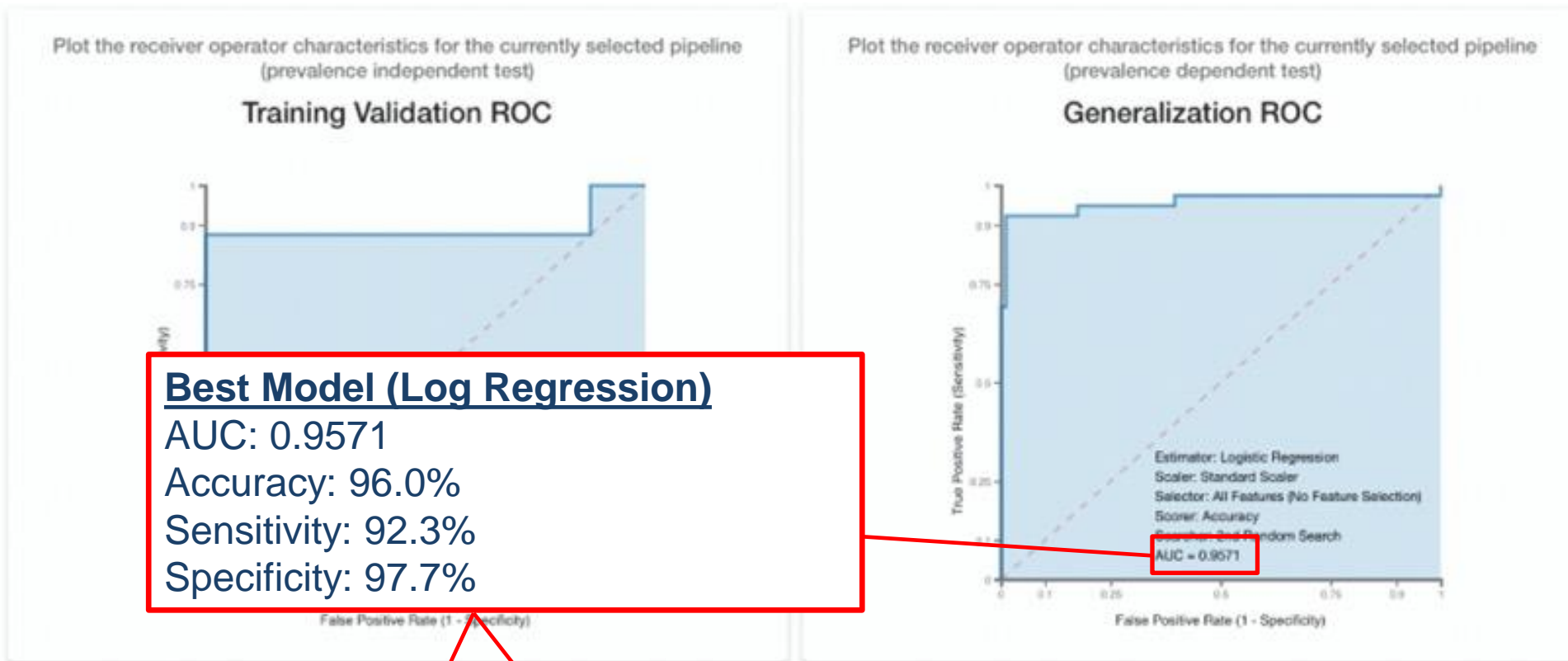
A



Estimator	$\overline{Sn+Sp}$ ↓	ROC AUC	Accuracy	F1	Sensitivity	Specificity	Brier Score	Scaler	Feature Selector	Scorer	Searcher
Logistic Regression	0.95	0.957	0.96	0.953	0.923	0.977	0.138	Standard Scaler	All Features (No Feature Selection)	Accuracy	2nd Random Search

AI/ML Enhanced Point-of-Care AKI Detection

A



Estimator	Sn+Sp ↓	ROC AUC	Accuracy	F1	Sensitivity	Specificity	Brier Score	Scaler	Feature Selector	Scorer	Searcher
Logistic Regression	0.95	0.957	0.96	0.95	0.923	0.977	0.138	Standard Scaler	All Features (No Feature Selection)	Accuracy	2nd Random Search

Emergence of Generative AI: Chatbots and Beyond



Role of AI Chatbots in Laboratory and Point-of-Care Testing



What NGAL values are predictive of AKI?



The specific NGAL values that are predictive of AKI may vary depending on the population being studied and the clinical context, and there is currently no universally agreed-upon threshold for NGAL levels to predict AKI. However, in general, higher NGAL values are associated with an increased risk of AKI.

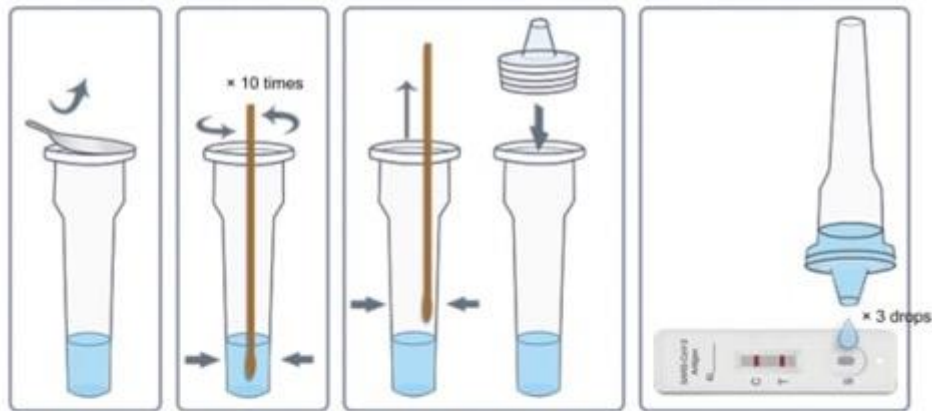


For example, a meta-analysis of studies on NGAL and AKI found that an NGAL value of 150 ng/mL or higher was associated with an increased risk of AKI, with higher values indicating a greater risk. However, it's important to note that this threshold may not be applicable in all clinical settings, and other factors such as patient age, comorbidities, and medications may also affect the predictive value of NGAL.

Ultimately, the interpretation of NGAL levels in the context of AKI risk should be made by a qualified healthcare provider who is familiar with the individual patient's medical history and clinical status.



Role of AI Chatbots in Laboratory and Point-of-Care Testing



NOTE: This figure is only used as a reference.

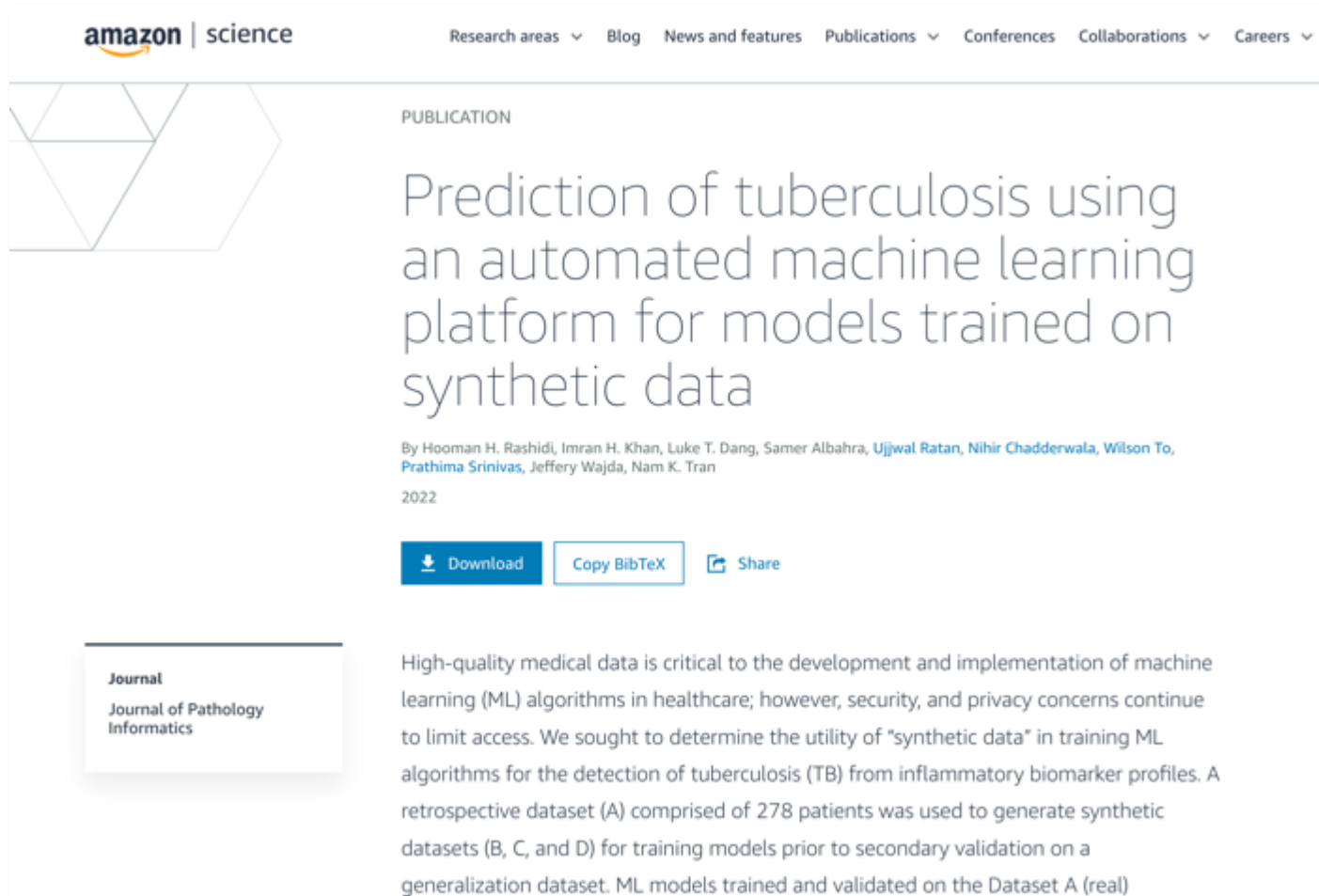
DA Tell me how to use a COVID antigen test

Here are the general steps for using a COVID-19 antigen test:

1. First, wash your hands thoroughly with soap and water, or use hand sanitizer.
2. Open the test kit and remove the contents. The kit should contain a test device, a buffer solution, a swab, and instructions.
3. Use the swab provided to collect a sample from the inside of your nose. Insert the swab into one nostril and rotate it several times, then repeat the process in the other nostril.
4. Dip the swab into the buffer solution provided and mix it thoroughly for a few seconds to ensure that the sample is well mixed.
5. Open the test device and add a few drops of the mixed sample to the test strip.
6. Wait for the recommended time, which is usually around 15-30 minutes, for the test result to appear.
7. Read the test result according to the manufacturer's instructions. A positive result typically indicates the presence of the COVID-19 antigen, while a negative result indicates the absence of the antigen.

It's important to note that the instructions may vary slightly depending on the specific test kit being used, so it's important to carefully read and follow the instructions provided with the kit. Additionally, while antigen tests can be useful in detecting COVID-19 infections, they are not as accurate as PCR tests and should not be relied upon as the sole method of diagnosis. If you have symptoms of COVID-19 or have been in close contact with someone who has tested positive, you should seek advice from a healthcare professional on the appropriate testing and management plan.

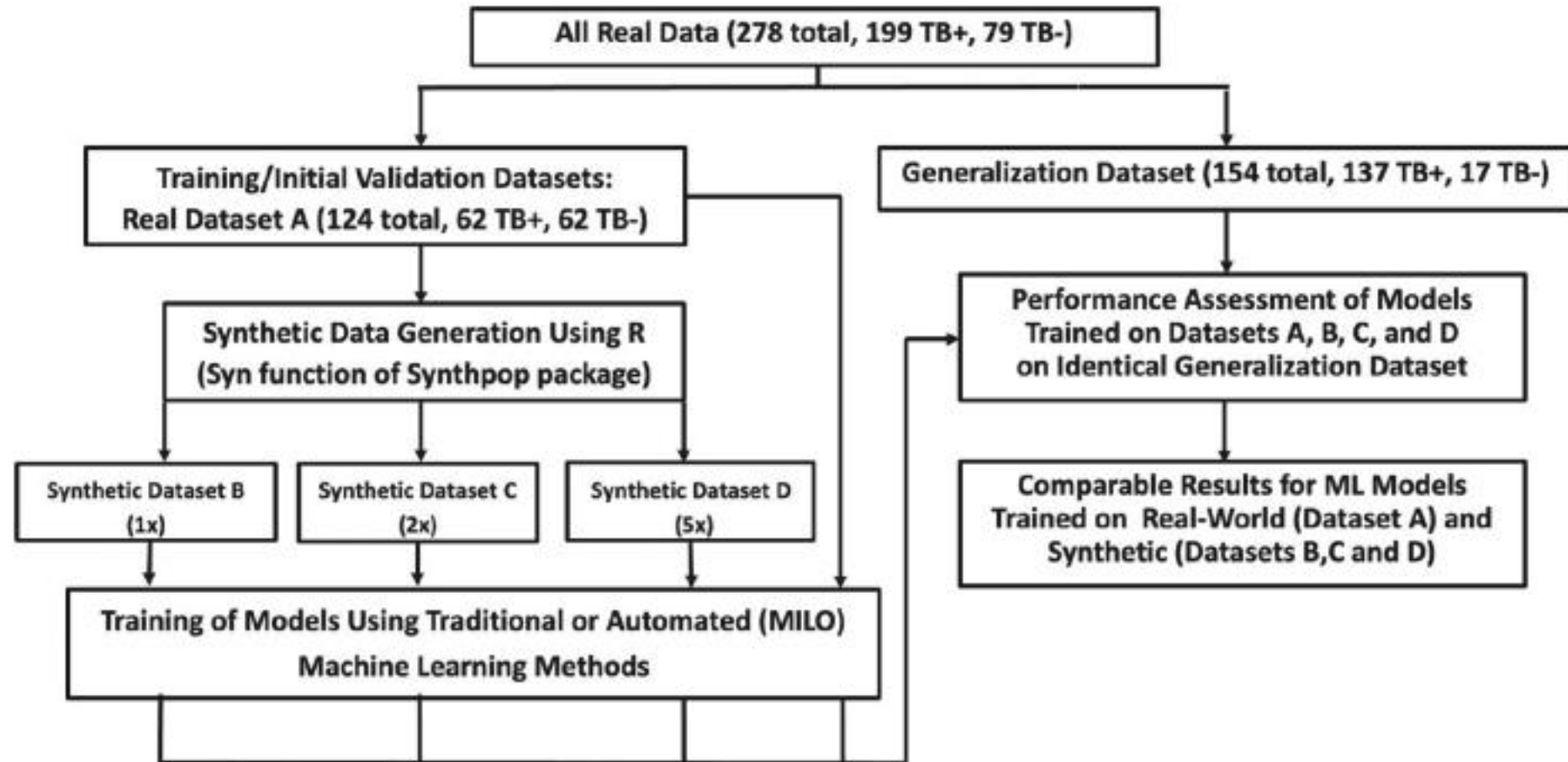
Generative AI for Producing Synthetic Data?



The screenshot shows the Amazon Science website with a navigation bar at the top containing 'amazon | science' and links for 'Research areas', 'Blog', 'News and features', 'Publications', 'Conferences', 'Collaborations', and 'Careers'. The main content area features a large title 'Prediction of tuberculosis using an automated machine learning platform for models trained on synthetic data' and a list of authors: 'By Hooman H. Rashidi, Imran H. Khan, Luke T. Dang, Samer Albahra, Ujjwal Ratan, Nihir Chadderwala, Wilson To, Prathima Srinivas, Jeffery Wajda, Nam K. Tran'. Below the authors is the year '2022' and three buttons: 'Download', 'Copy BibTeX', and 'Share'. A 'Journal' box on the left identifies the publication as 'Journal of Pathology Informatics'. The abstract text reads: 'High-quality medical data is critical to the development and implementation of machine learning (ML) algorithms in healthcare; however, security, and privacy concerns continue to limit access. We sought to determine the utility of "synthetic data" in training ML algorithms for the detection of tuberculosis (TB) from inflammatory biomarker profiles. A retrospective dataset (A) comprised of 278 patients was used to generate synthetic datasets (B, C, and D) for training models prior to secondary validation on a generalization dataset. ML models trained and validated on the Dataset A (real)

- Synthetic data is a rapidly evolving field whereby AI is used to produce unique data based on real clinical data.
- This helps increase sample size, but also reduces upfront barriers in accessing clinical data (e.g., IRB, time, labor).
- Synthetic data is of course not perfect, but serves as a starting point.
- Likewise, synthetic data could be used to help with AI/ML quality control and perhaps serve as a form of proficiency testing.

Use of Synthetic Data for Developing AI/ML Algorithms to Predict Tuberculosis



Use of Synthetic Data for Developing AI/ML Algorithms to Predict Tuberculosis

Model performances based on the “real” secondary dataset	Trained on dataset A real data (95% CI)	Trained on dataset B (synthetic data ×1) (95% CI)	Trained on dataset C (synthetic data ×2) (95% CI)	Trained on dataset D (synthetic data ×5) (95% CI)
MILO’s best models	MILO GBM	MILO SVM	MILO DNN	MILO DNN
ROC-AUC	0.95 (0.87–1)	0.83 (0.63–1)	0.91 (0.8–1)	0.55 (0.48–0.62)
Accuracy	90 (84–95)	91 (85–95)	71 (63–78)	54 (46–62)
Sensitivity	89 (83–94)	93 (87–96)	67 (59–75)	49 (40–58)
Specificity	100 (81–100)	77 (50–93)	100 (81–100)	94 (71–99)
MILO’s best RF models	MILO RF	MILO RF	MILO RF	MILO RF
ROC-AUC	0.96 (0.82–1)	0.77 (0.67–0.87)	0.87 (0.77–0.97)	0.66 (0.52–0.8)
Accuracy	89 (83–93)	71 (63–78)	74 (66–81)	56 (48–64)
Sensitivity	88 (81–93)	69 (60–76)	72 (64–80)	53 (44–61)
Specificity	100 (81–100)	88 (64–99)	88 (64–99)	82 (57–96)
Non-MILO RF models	Non-MILO RF	Non-MILO RF	Non-MILO RF	Non-MILO RF
ROC-AUC	0.97 (0.94–1)	0.73 (0.60–0.88)	0.83 (0.71–0.92)	0.68 (0.57–0.82)
Accuracy	77 (70–84)	62 (54–69)	64 (56–72)	39 (31–47)
Sensitivity	75 (66–82)	61 (52–69)	64 (55–72)	40 (32–49)
Specificity	100 (81–100)	71 (44–90)	71 (44–90)	29 (10–56)

DNN = deep neural network, GBM = gradient boosting machine, RF = random forest, SVM = support vector machine.

Use of Synthetic Data for Developing AI/ML Algorithms to Predict Tuberculosis

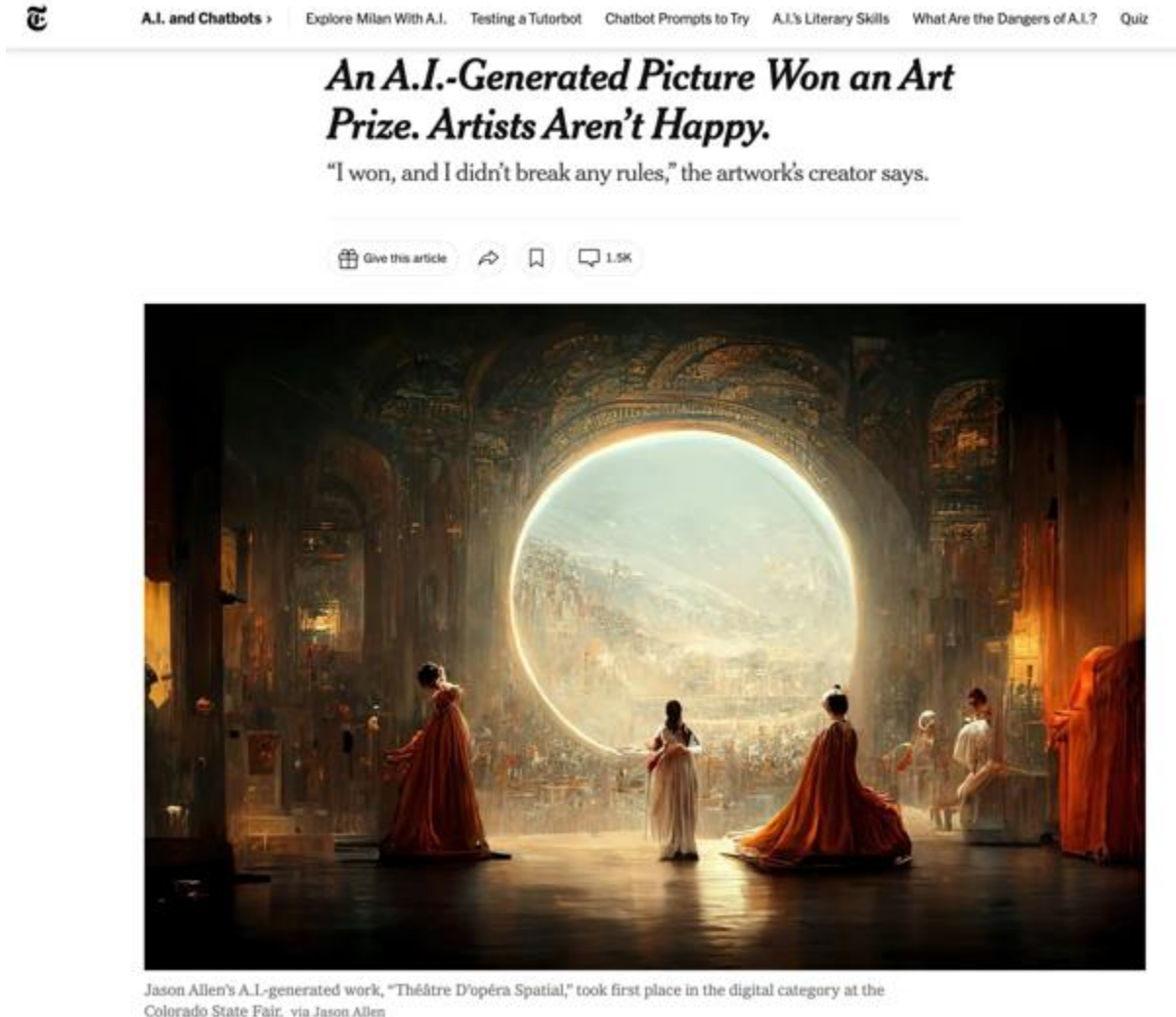
Model performances based on the “real” secondary dataset	Trained on dataset A real data (95% CI)	Trained on dataset B (synthetic data ×1) (95% CI)	Trained on dataset C (synthetic data ×2) (95% CI)	Trained on dataset D (synthetic data ×5) (95% CI)
MILO’s best models	MILO GBM	MILO SVM	MILO DNN	MILO DNN
ROC-AUC	0.95 (0.87–1)	0.83 (0.63–1)	0.91 (0.8–1)	0.55 (0.48–0.62)
Accuracy	90 (84–95)	91 (85–95)	71 (63–78)	54 (46–62)
Sensitivity	89 (83–94)	93 (87–96)	67 (59–75)	49 (40–58)
Specificity	100 (81–100)	77 (58–93)	100 (81–100)	94 (71–99)
MILO’s best RF models	MILO RF	MILO RF	MILO RF	MILO RF
ROC-AUC	0.96 (0.82–1)	0.77 (0.67–0.87)	0.87 (0.77–0.97)	0.66 (0.52–0.8)
Accuracy	89 (83–93)	71 (63–78)	74 (66–81)	56 (48–64)
Sensitivity	88 (81–93)	69 (60–76)	72 (64–80)	53 (44–61)
			88 (64–99)	82 (57–96)
			Non-MILO RF	Non-MILO RF
			0.83 (0.71–0.92)	0.68 (0.57–0.82)
			64 (56–72)	39 (31–47)
			64 (55–72)	40 (32–49)
			71 (44–90)	29 (10–56)

Result showed that AI/ML models still exhibit reduced performance when trained only on synthetic data, however, performance also varies by AI/ML methods used. Using the MILO automated ML platform, deep neural network (DNN) using dataset C provided the least decrease in performance compared to real world data.

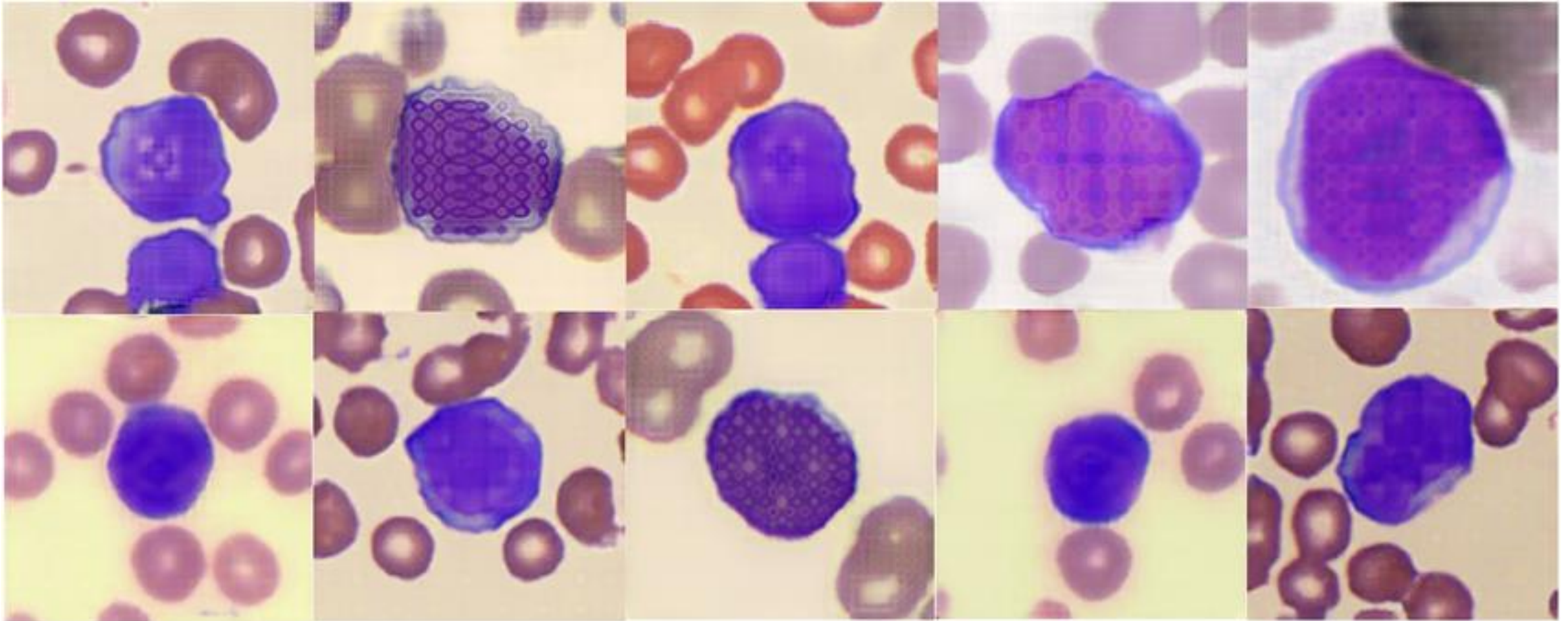
DNN = deep neural network, GBM = gradient boosting machine, RF = random forest, SVM = support vector machine.

Generative AI to Produce Medical Images?

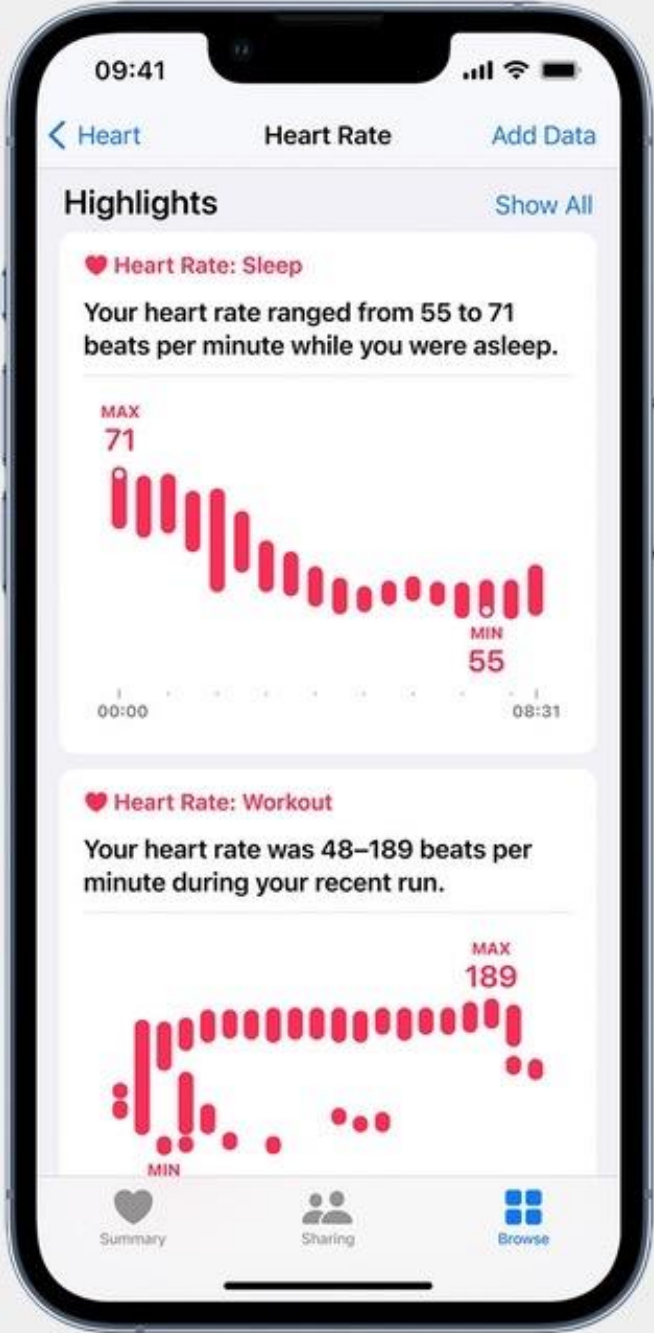
- Generative AI systems have been used to produce photo-realistic images.
- Images have been able to replicate real-world objects and people.
- Generative AI artwork has also competed and won against human artists.
- However, the use of generative AI images in healthcare is still evolving.



Artificial Intelligence Generated Leukemia (Blast) Cells – Work in Progress



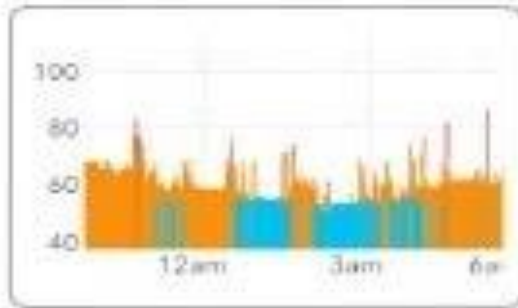
DIGITAL POC BIOMARKERS?



Sharable Home Testing Data?

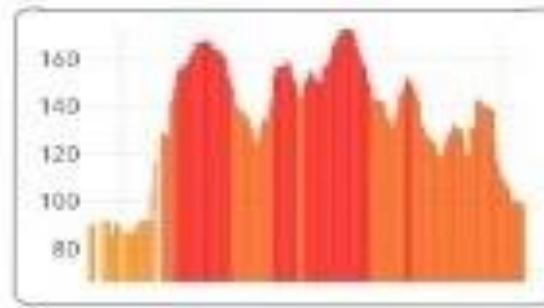
Sleep

Restful Sleep



Fitness

Running



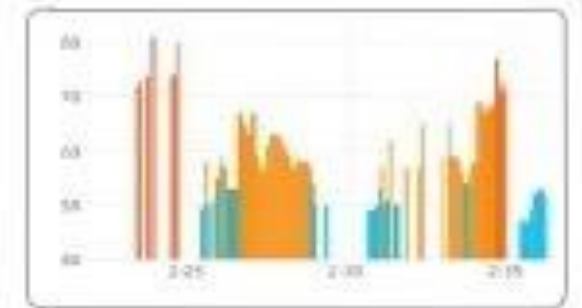
Stress

Driving in Rush Hour

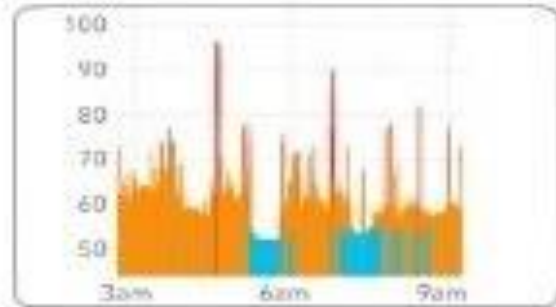


Illness

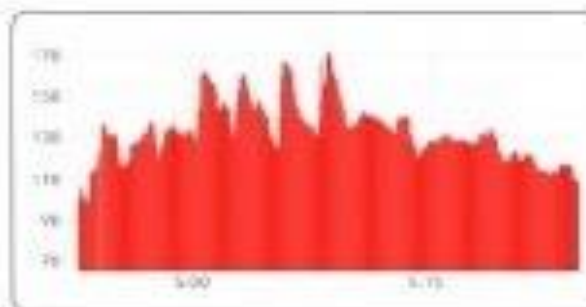
Atrial Fibrillation



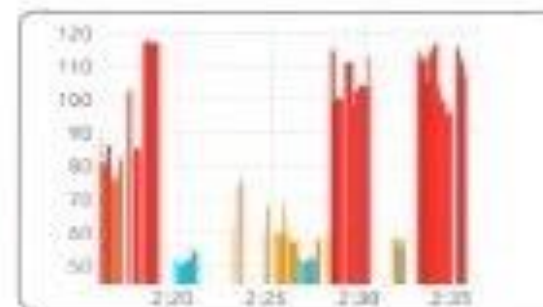
Sleep after Alcohol



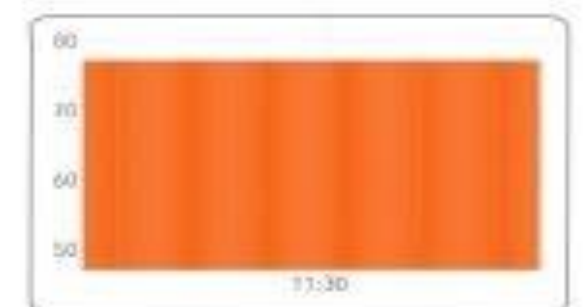
Hockey



Microsoft interview



Atrial Flutter



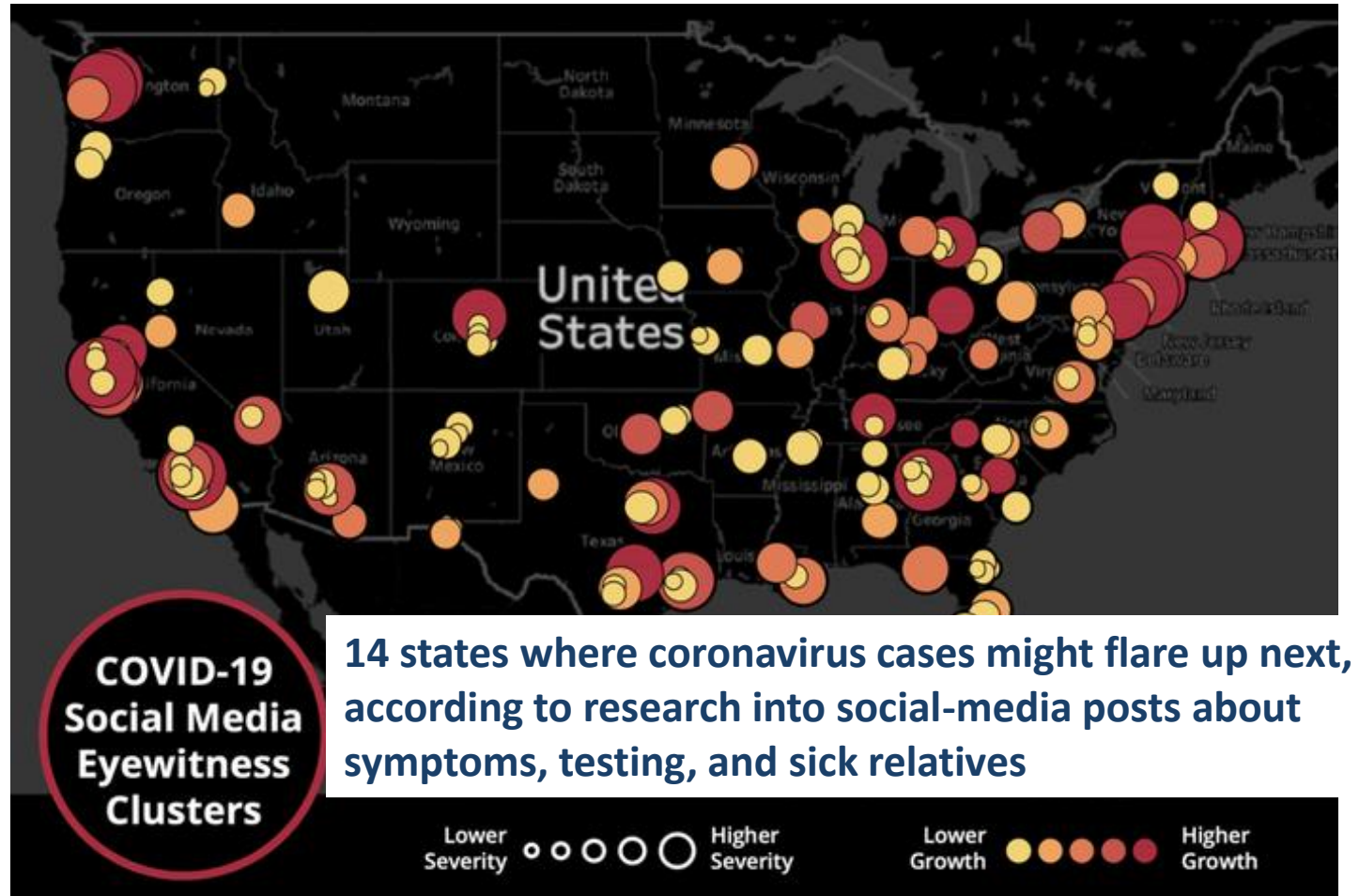
Sharable Home Testing Data?

- Wearable and home monitoring devices have become powerful health data generating tools.
- These include existing smart watches and phones, but also smart rings, and sleep monitoring systems that aggregate real-time health data.
- Such data could be coupled to other testing modalities and AI/ML to predict a range of diseases.

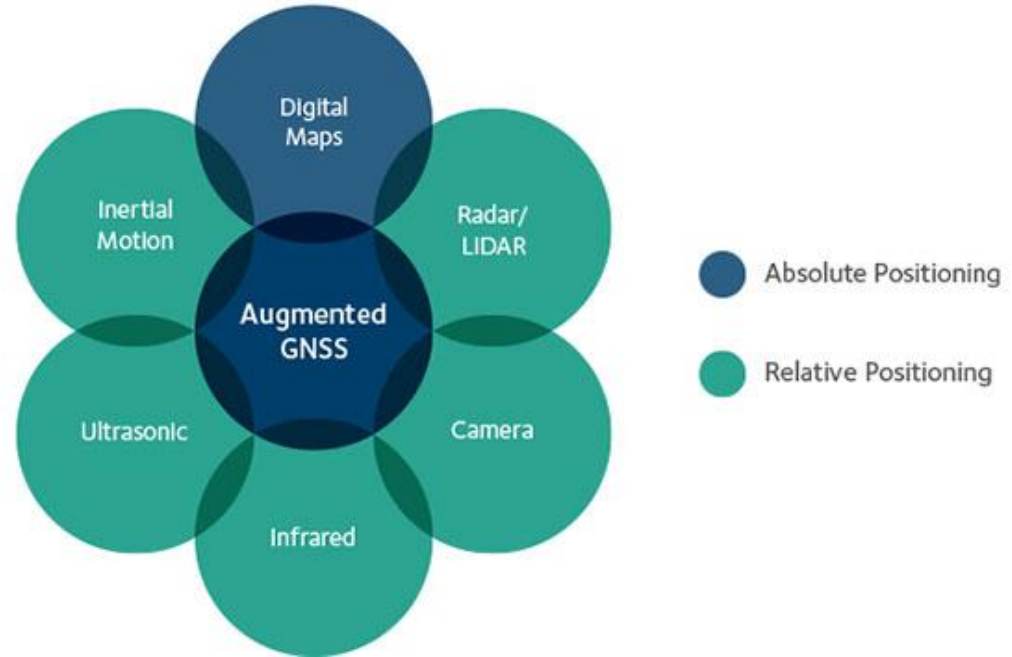
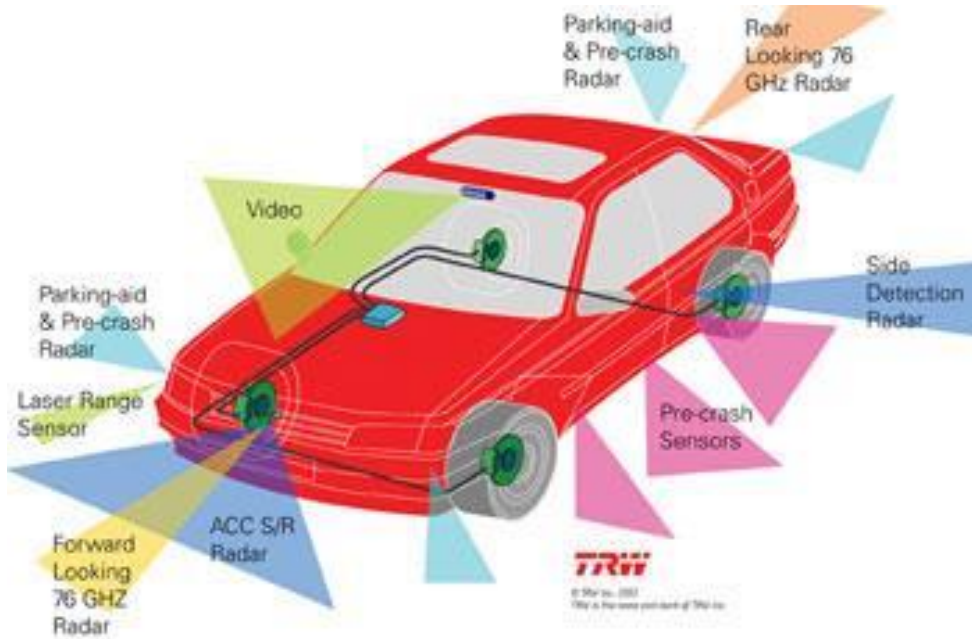


Open Source Intelligence → Catching Disasters Early?

- Shared data from POC devices and even social media could be used for population health.
- During the COVID-19 pandemic, social media posts could predict where the next outbreak will occur.
- Likewise, IVD manufacturers were already using crowd-sourced instrument data to predict out COVID-19 outbreaks based on Ct-values.
- Same could be applied to de-identified home or DTC data.

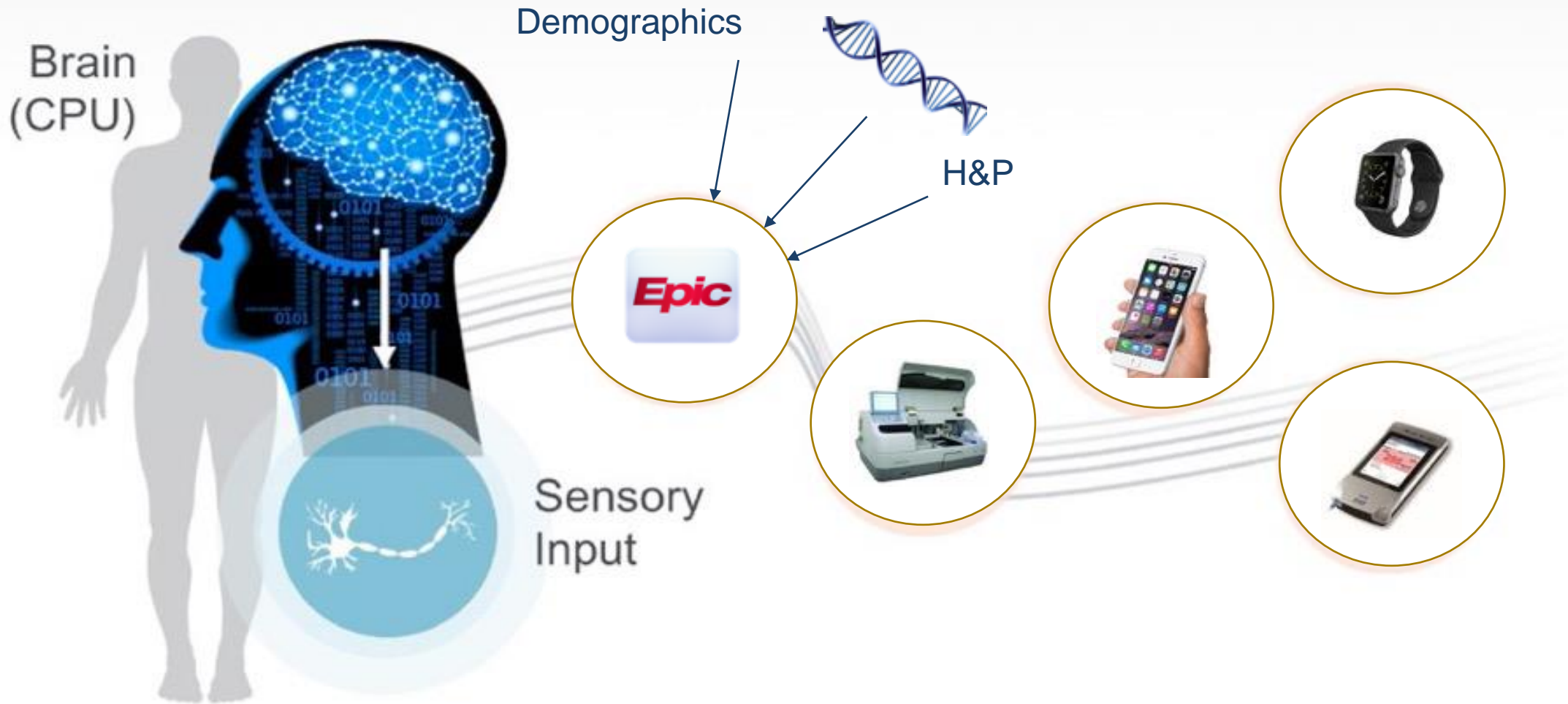


Future of Medical Care: Sensor Fusion



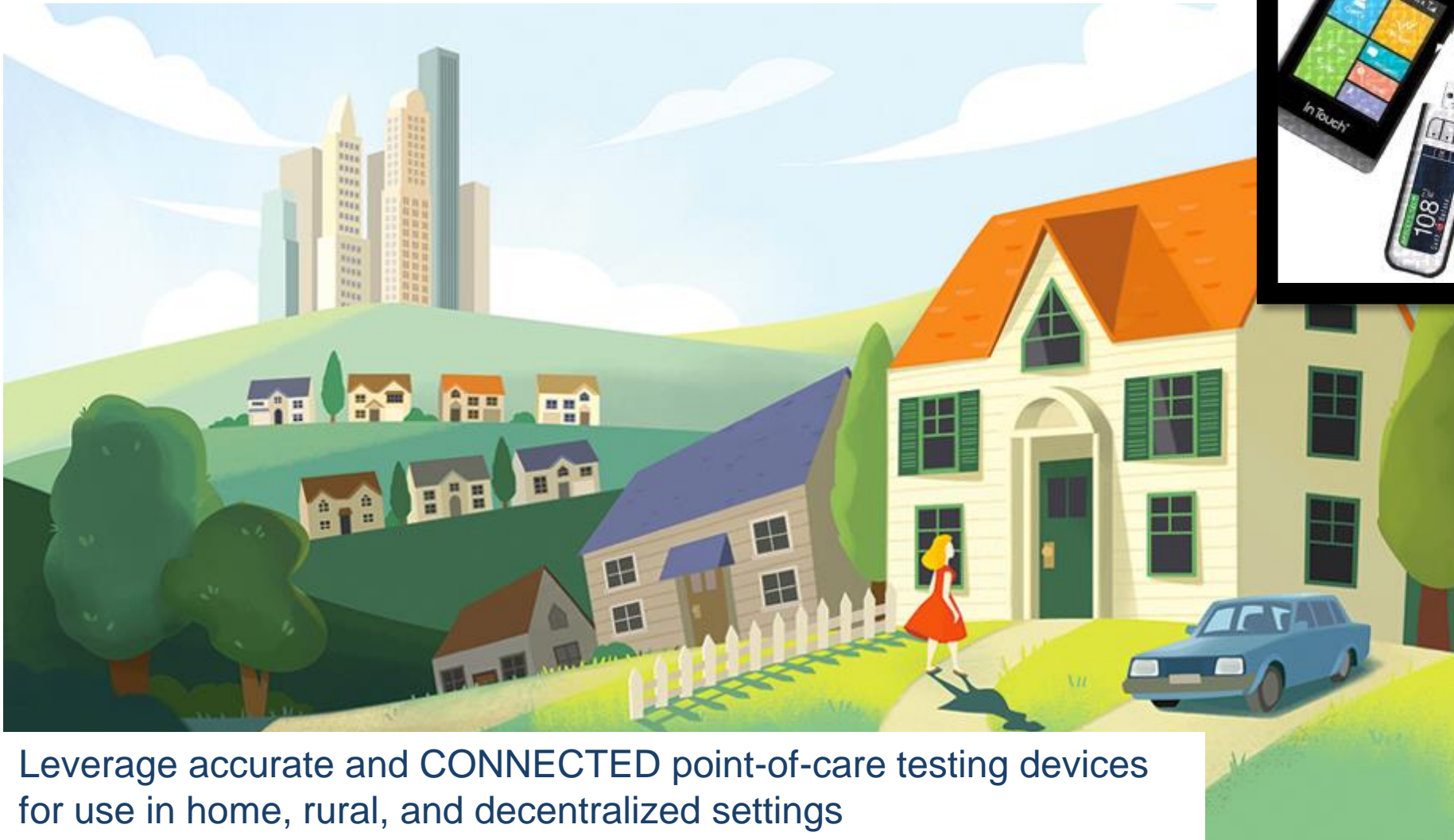
Sensor fusion is a process by which data from several different **sensors** are "fused" to compute something more than could be determined by any one **sensor** alone.

“Medical” Sensor Fusion



Medical sensor future may be the future of patient care. Integration of multiple sources of medical information (POCT, lab, smart devices, genetic testing, EHR data) into meaningful and actionable results.

Future Vision of Hybrid Automated Laboratory Testing



Leverage accurate and CONNECTED point-of-care testing devices for use in home, rural, and decentralized settings

Future Vision of Hybrid Automated Laboratory Testing

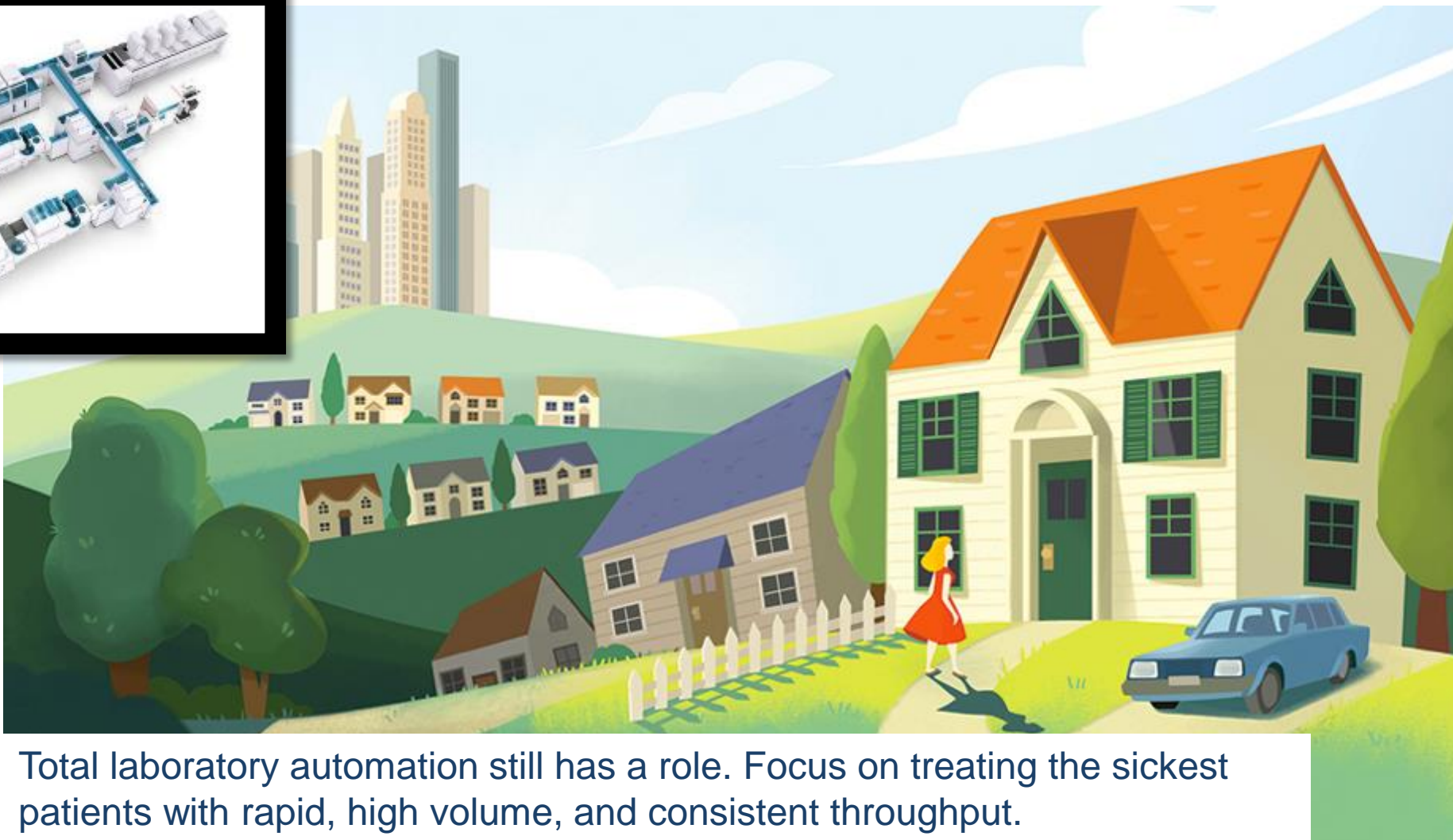


Future Vision of Hybrid Automated Laboratory Testing

POCT no manifests in multiple forms including wearable smart devices, molecular tests, and multiplex platforms.

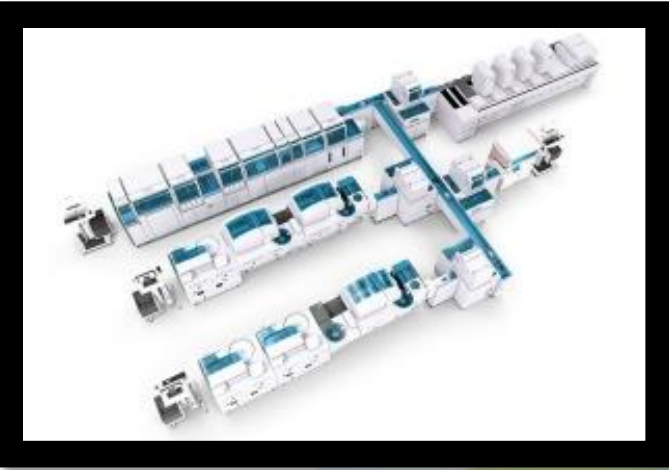


Future Vision of Hybrid Automated Laboratory Testing

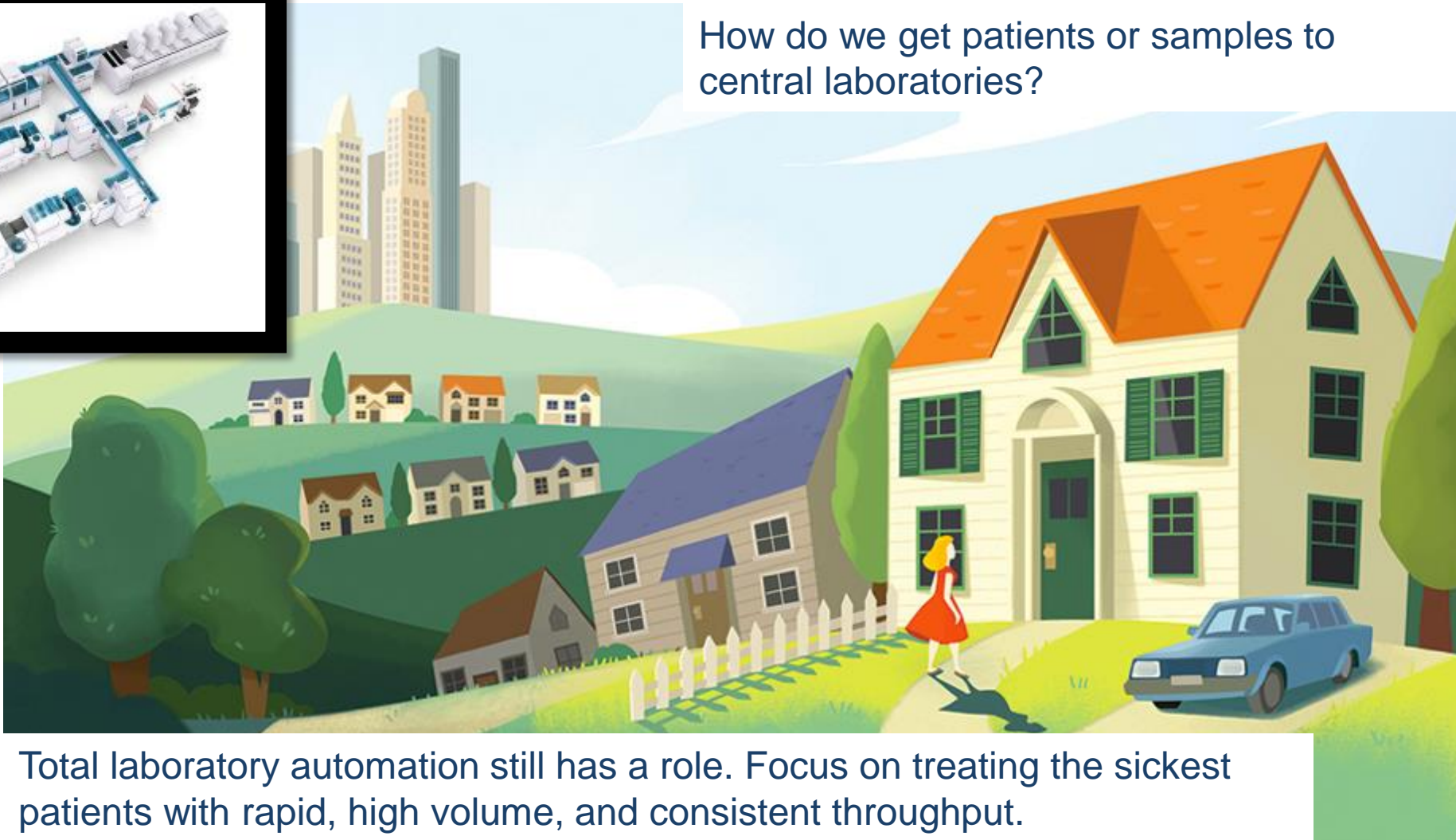


Total laboratory automation still has a role. Focus on treating the sickest patients with rapid, high volume, and consistent throughput.

Future Vision of Hybrid Automated Laboratory Testing



How do we get patients or samples to central laboratories?



Total laboratory automation still has a role. Focus on treating the sickest patients with rapid, high volume, and consistent throughput.

Future Vision of Hybrid Automated Laboratory Testing



Employ disruptive (automated) technologies such as drone aircraft. Speed and cost savings by removing patient travel and couriers.



Future Vision of Hybrid Automated Laboratory Testing



...and self driving cars

Summary

- Artificial intelligence and machine learning is gaining traction in healthcare including laboratory applications.
- Prior applications revolve around image recognition for blood/fluid image analysis, with more recent uses for disease recognition from electronic health record data.
- Exciting areas of development for AI/ML includes the use for genomic data analysis, spectral analysis, and clinical decision support for point-of-care testing.
- Synthetic data is also an exciting developing area, but has much room to improve.
- Generative AI with the use of chatbots is a very recent development that could be used to help with clinical decision making, as well as operator training/education.
- Future of AI and ML in the laboratory and POCT settings will focus on medical sensor fusion – integrating all sources of data to support testing operations.

Questions?