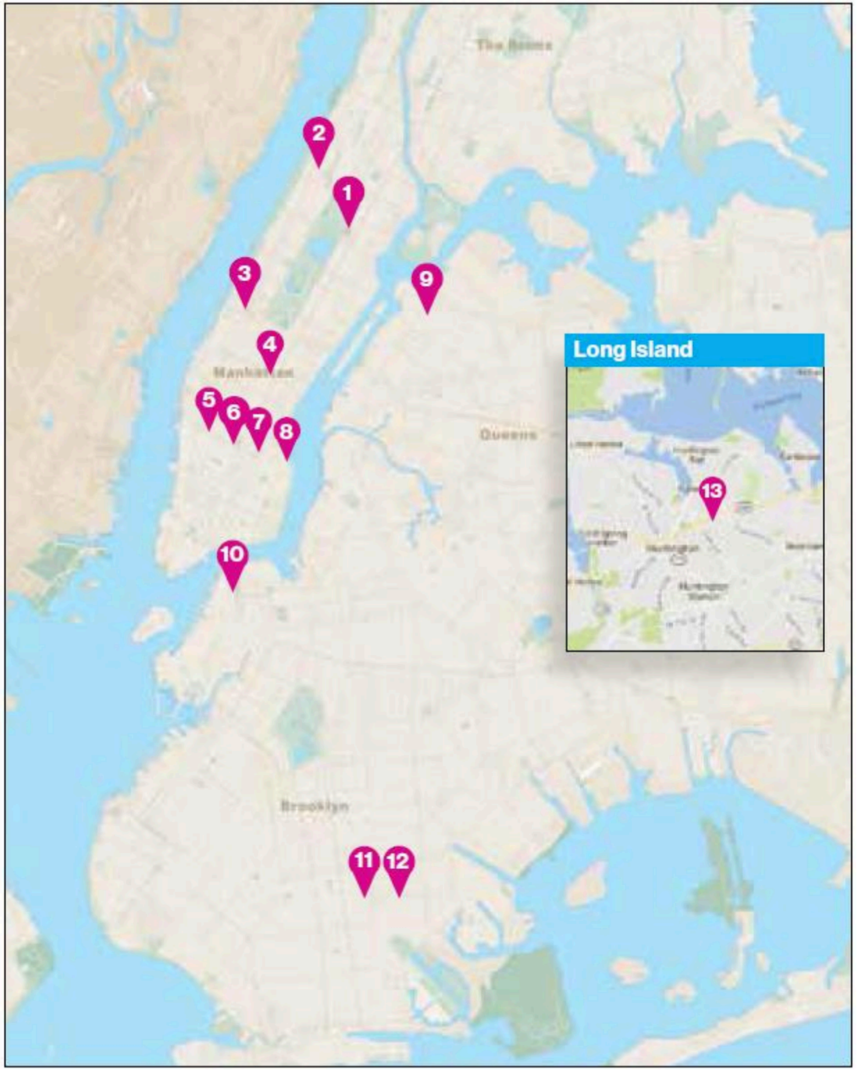
COVID-19 and the MSHS

THE MOUNT SINAI COVID INFORMATICS CENTER (MSCIC)

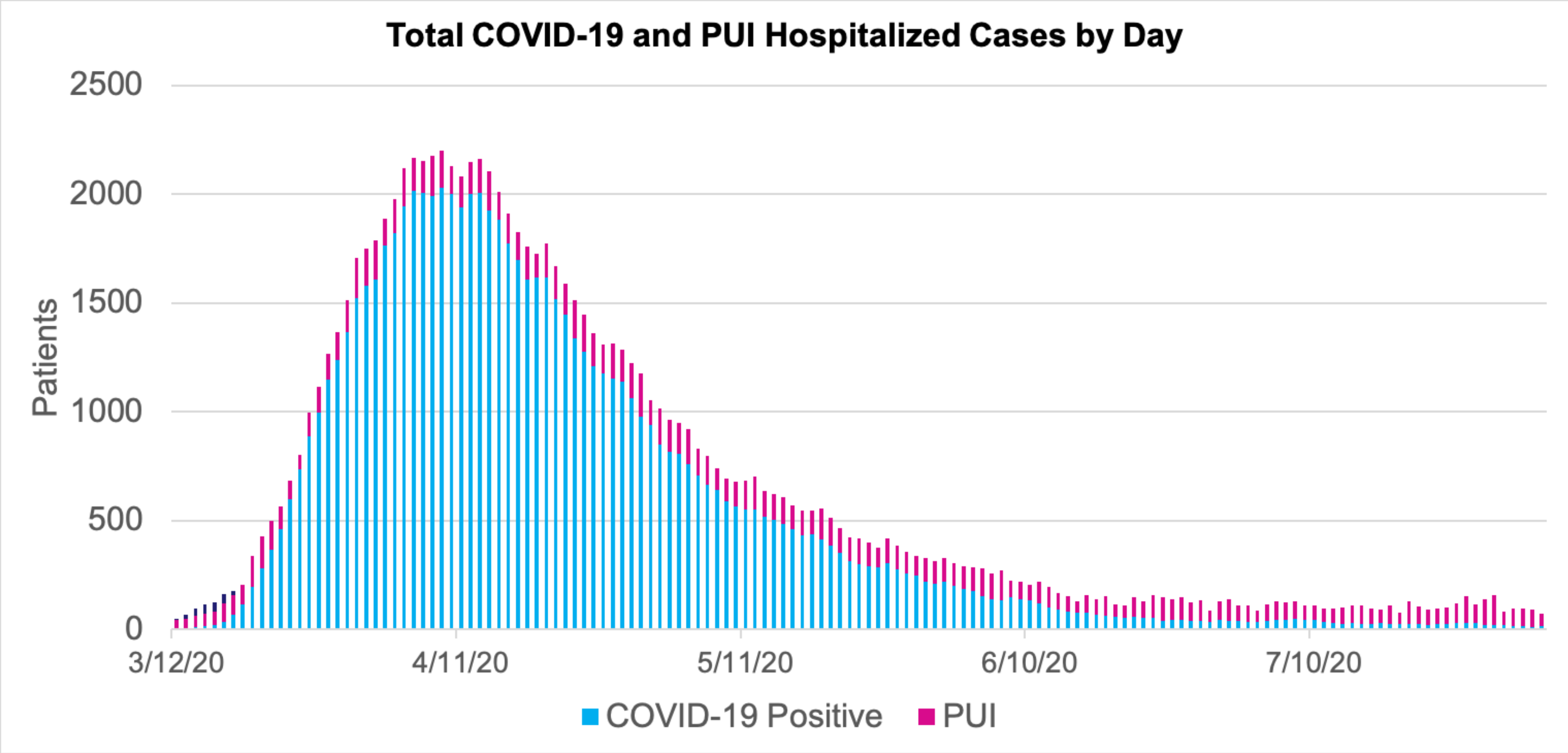
Girish N Nadkarni, MD, MPH

Icahn School of Medicine at Mount Sinai

The Mount Sinai Health System (MSHS)



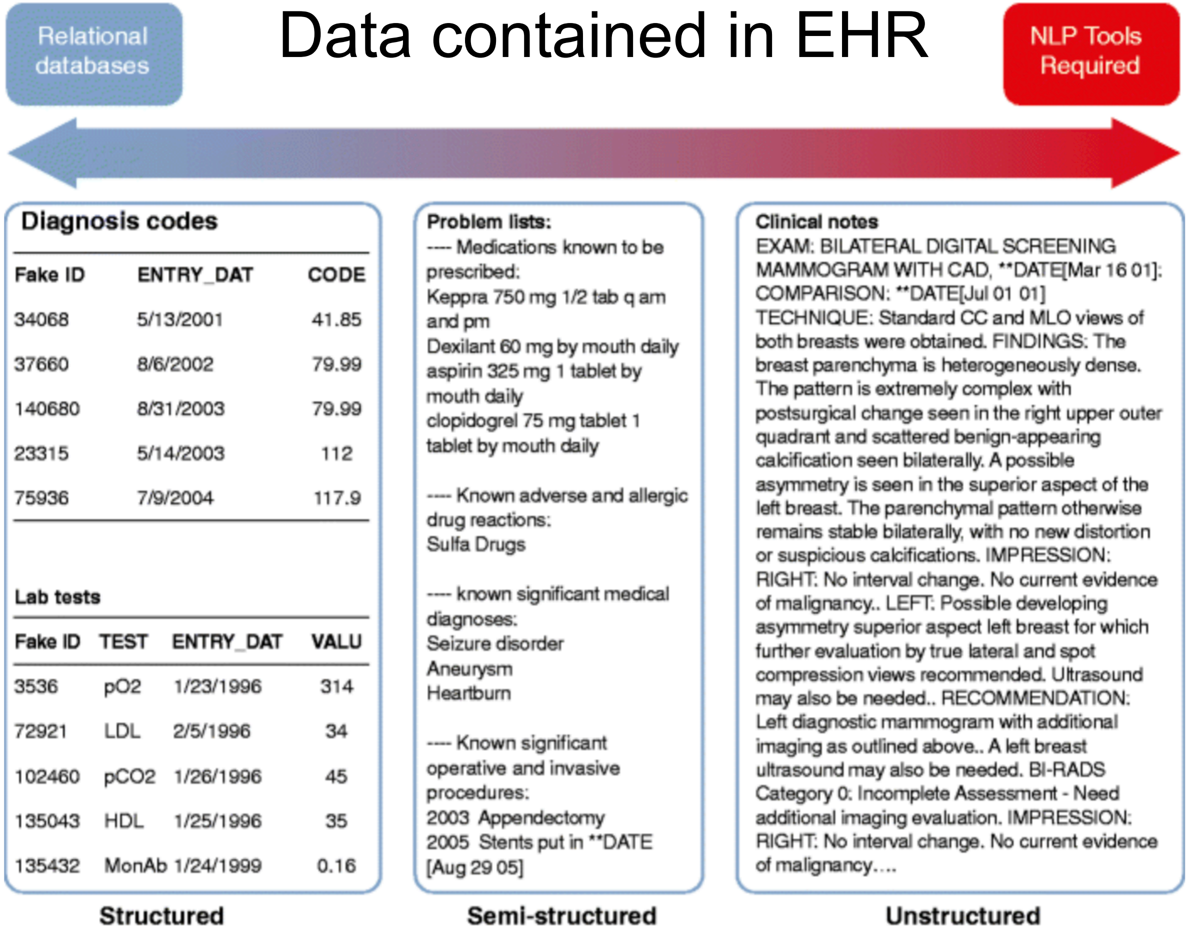
The Epicenter of the COVID-19 Pandemic in the United States




BIG DATA
IN A TIME OF
BIG VIRUSES.
How the World is Using Data &
Analytics to Fight COVID-19

Data contained in EHR

- Demographics
- Medical history
- Vital signs
- Diagnoses
- Medications
- Treatment plans
- Immunization history
- Radiology images
- Laboratory results



Wei-Qi, W. & Denny, J.C. Genome Medicine, 2015.

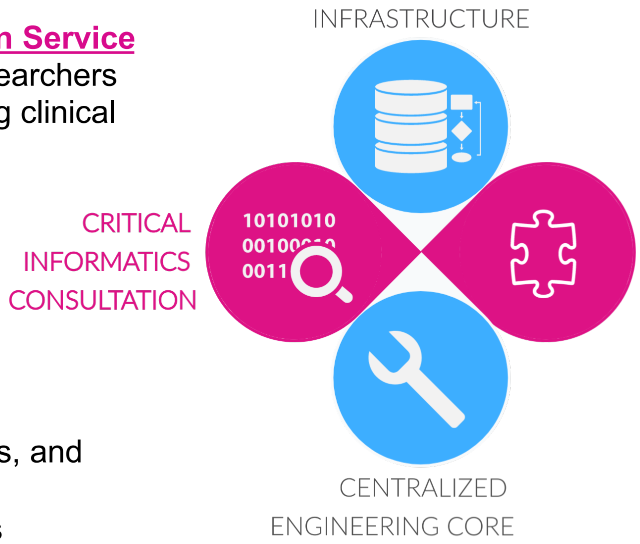
Mount Sinai COVID Informatics Center

Fighting Covid-19 with the Power of Data

Informatics Crisis Response Platform

Critical Informatics Consultation Service provides MSHS clinicians and researchers easy-to-digest answers to pressing clinical questions within 24 hours

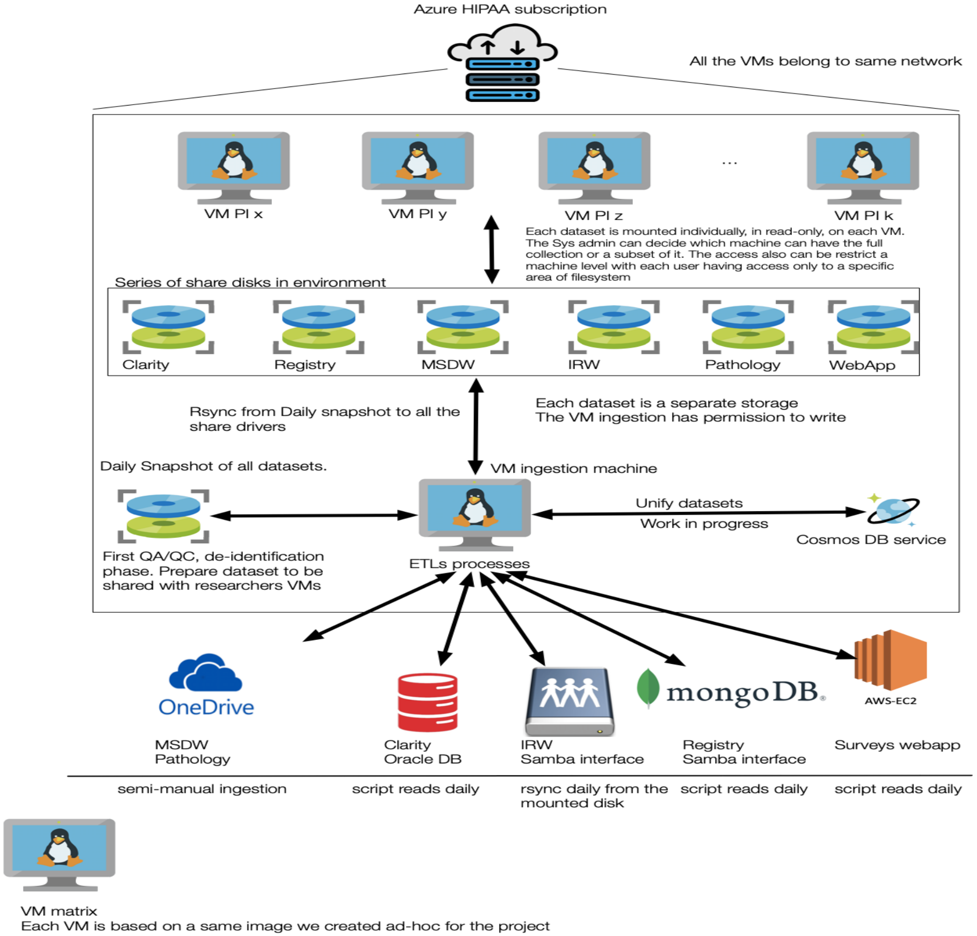
Centralized Engineering Core A team of highly trained computer scientists, engineers, informaticists, and researchers who are dedicated to productizing this Informatics Crisis Response Platform to enable MSHS to be battle ready in this and in future crises



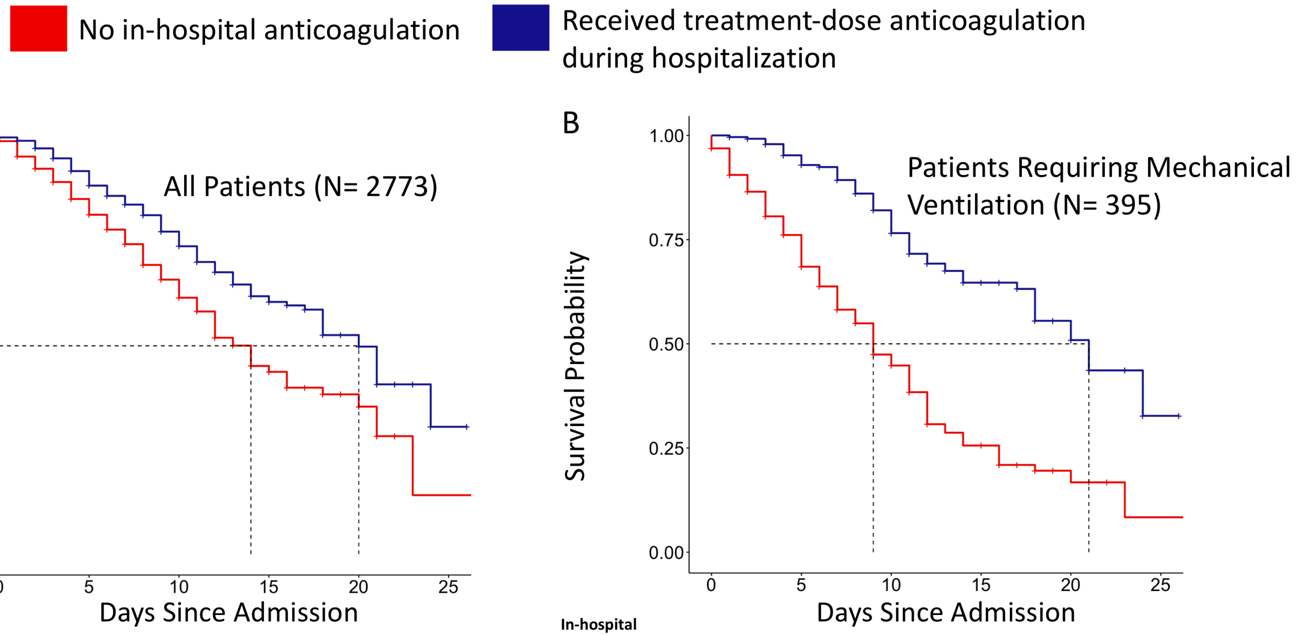
Infrastructure supported by Microsoft Azure cloud computing services, MSCIC has built and maintains a ground truth harmonized dataset that integrates data streams from MSHS clinical data (e.g. EHR, Imaging, Pathology) along with novel research data sets (e.g. –omics, digital health tracking, immune biomarkers)

Rapid Clinical Intervention Toolkit facilitates the practice of evidence-based medicine in the MSHS by feeding insights from data science into the daily workflow via the electronic medical record

MSCIC Data Platform



ASSOCIATION OF ANTICOAGULATION WITH MORTALITY

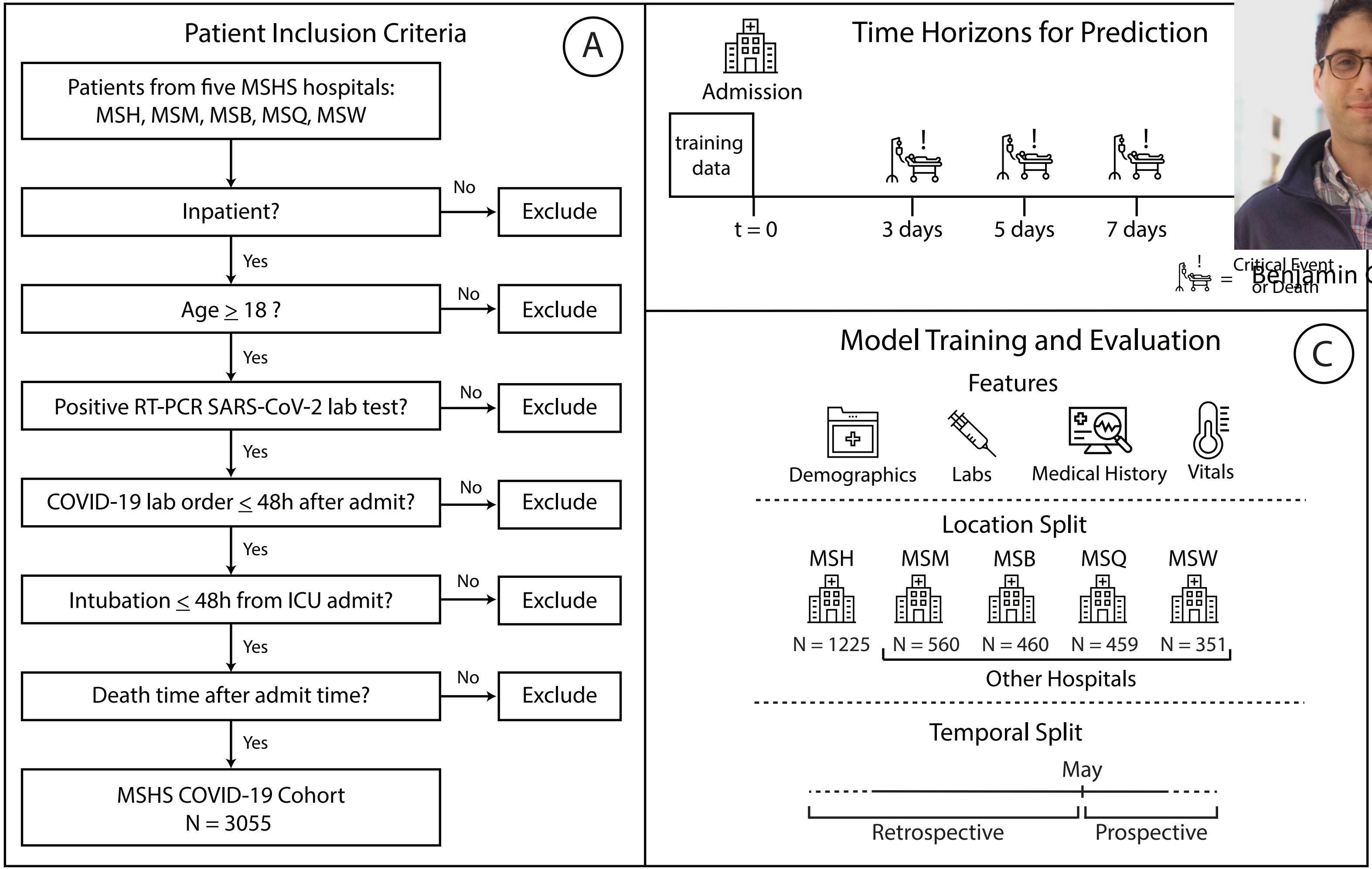


In-hospital Anticoagulation — Number at Risk (A)

Yes	786	538	266	90	19	3
No	1987	977	296	71	13	1

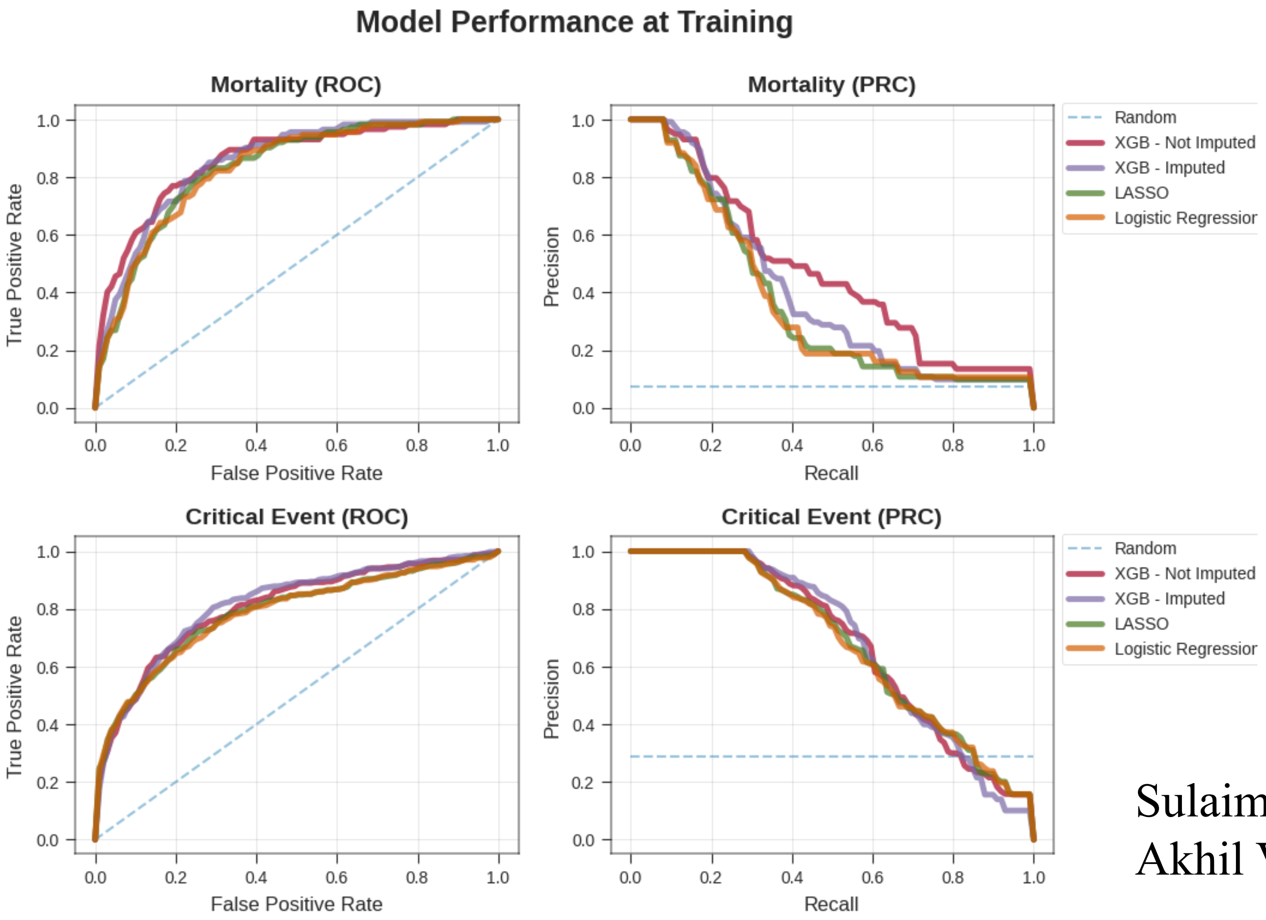
In-hospital Anticoagulation — Number at Risk (B)

Yes	234	197	137	65	14	3
No	161	100	54	25	7	1



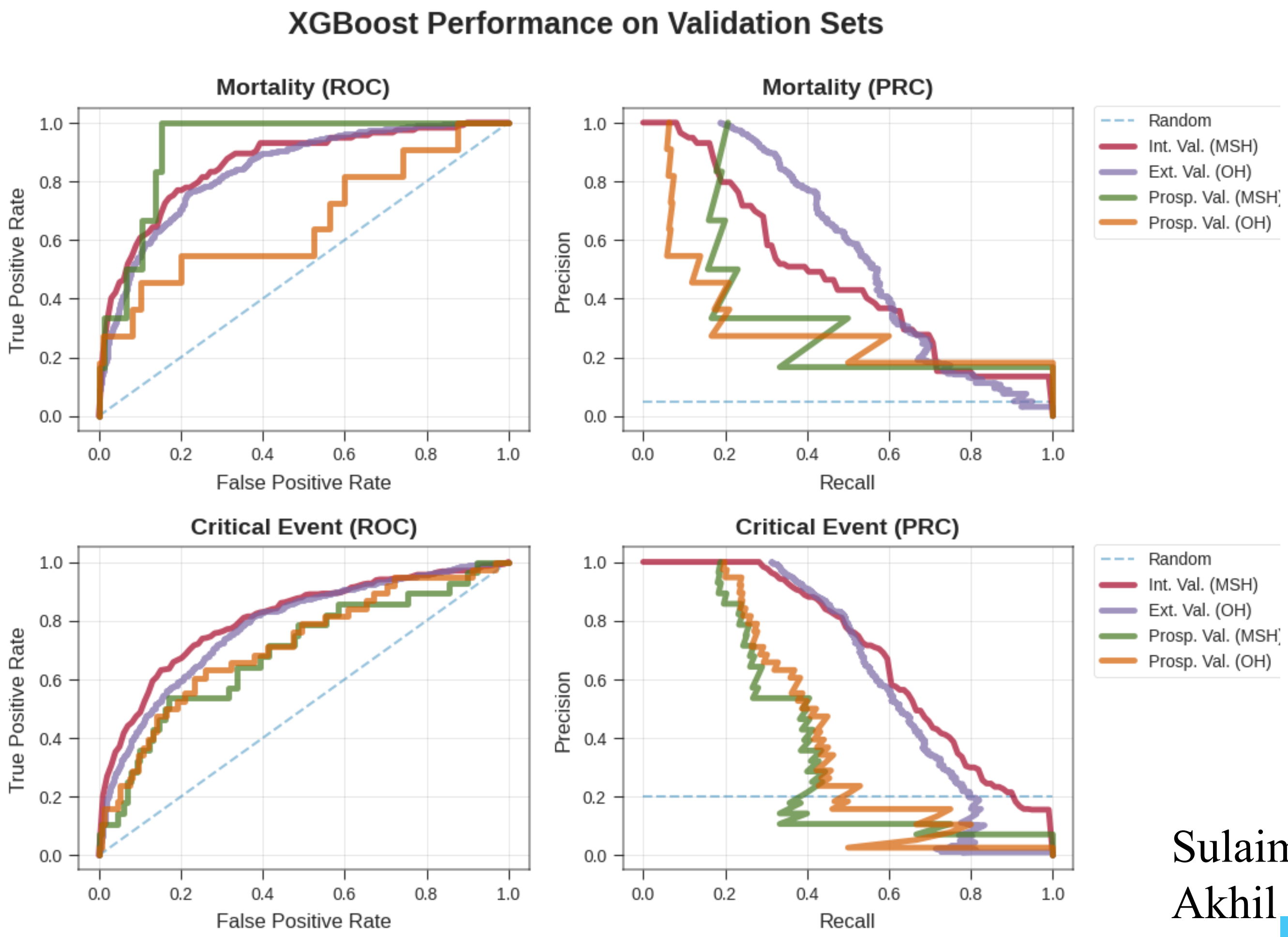

Benjamin Glicksberg

Performance at MSH (train + CV)



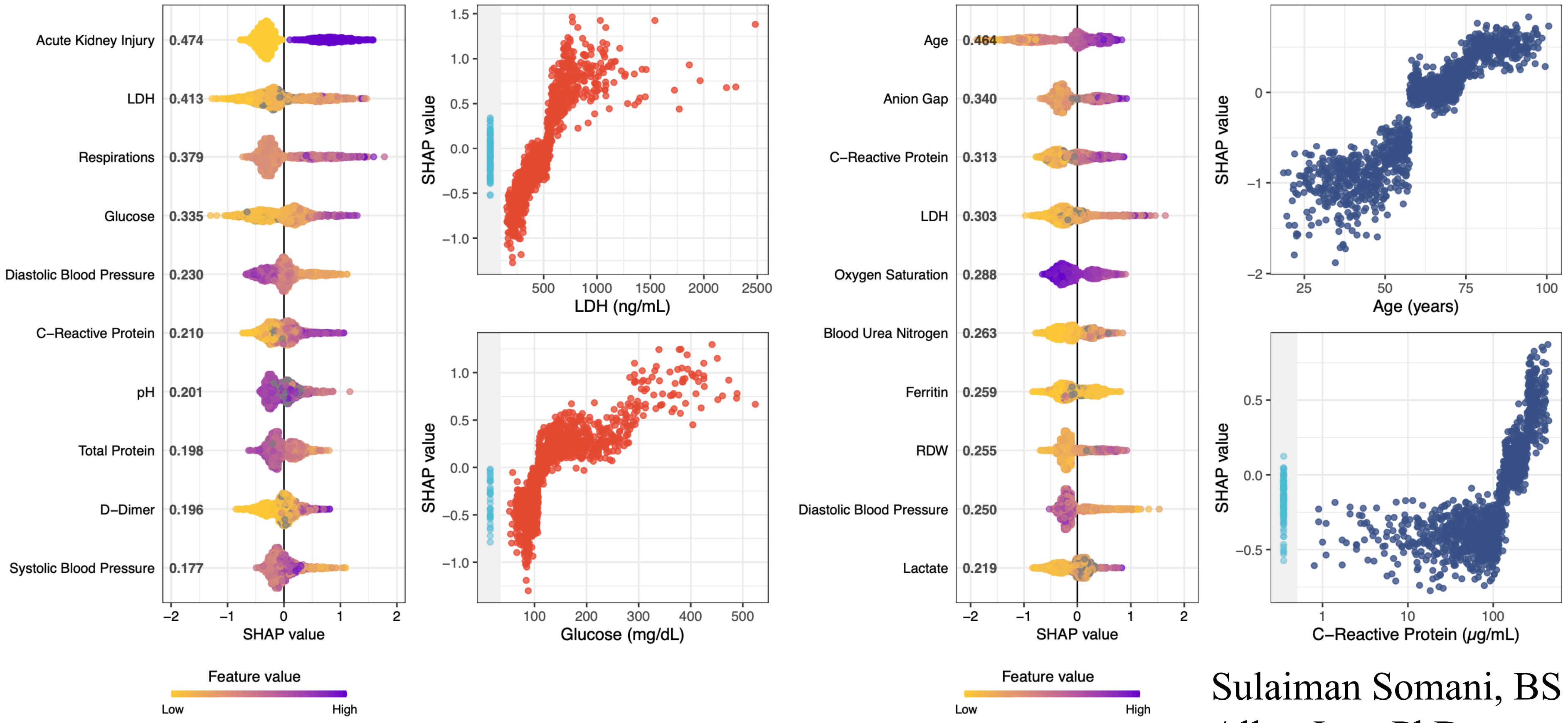
Sulaiman Somani, BS
Akhil Vaid, MD

External and temporal performance (validation)



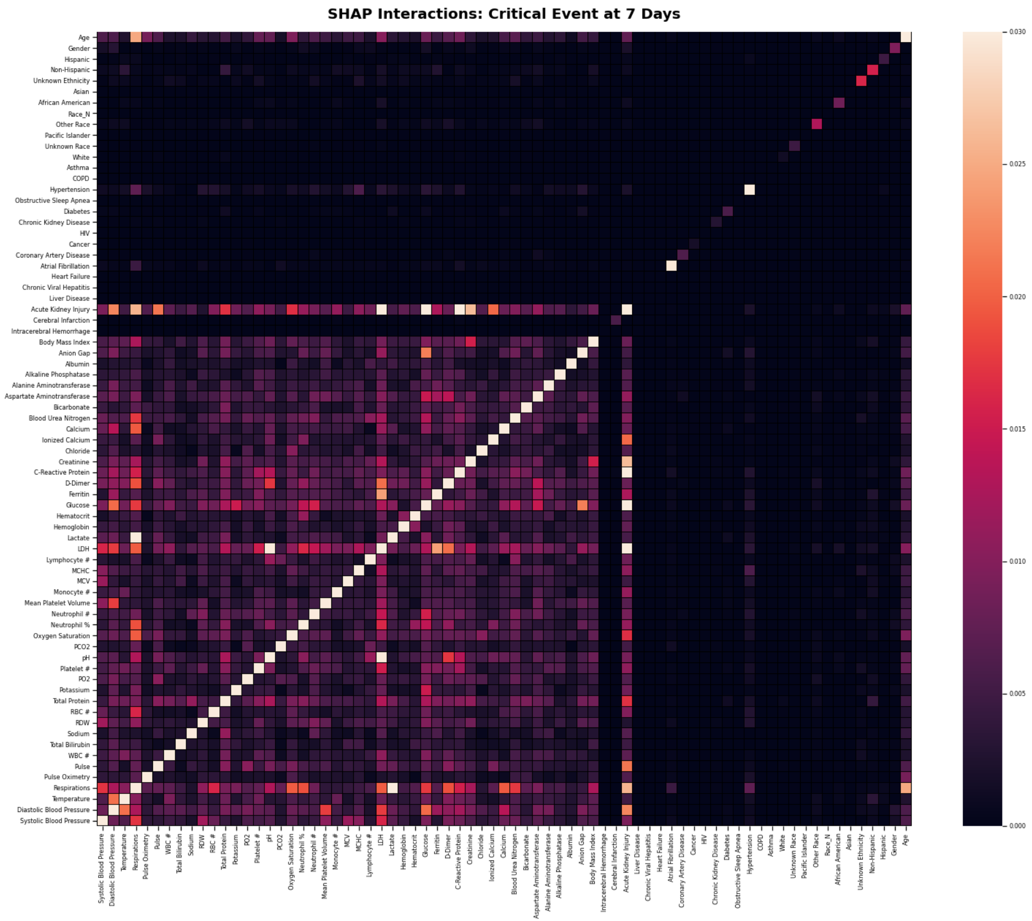
Sulaiman Somani, BS
Akhil Vaid, MD

What did the model learn?

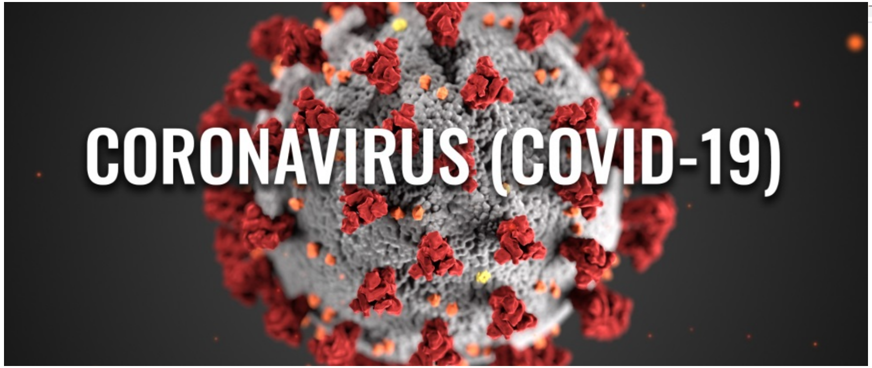


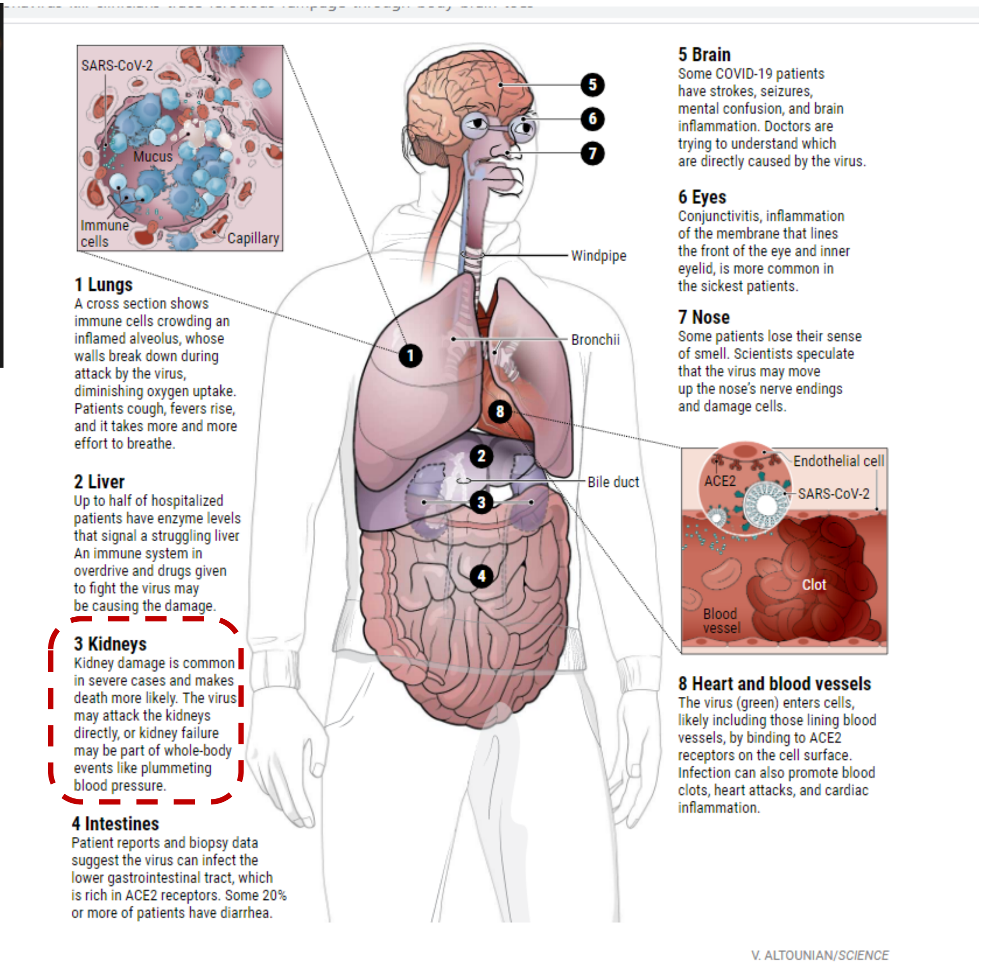
Sulaiman Somani, BS
Allan Just, PhD

Interactions between features can further reveal what the model learned



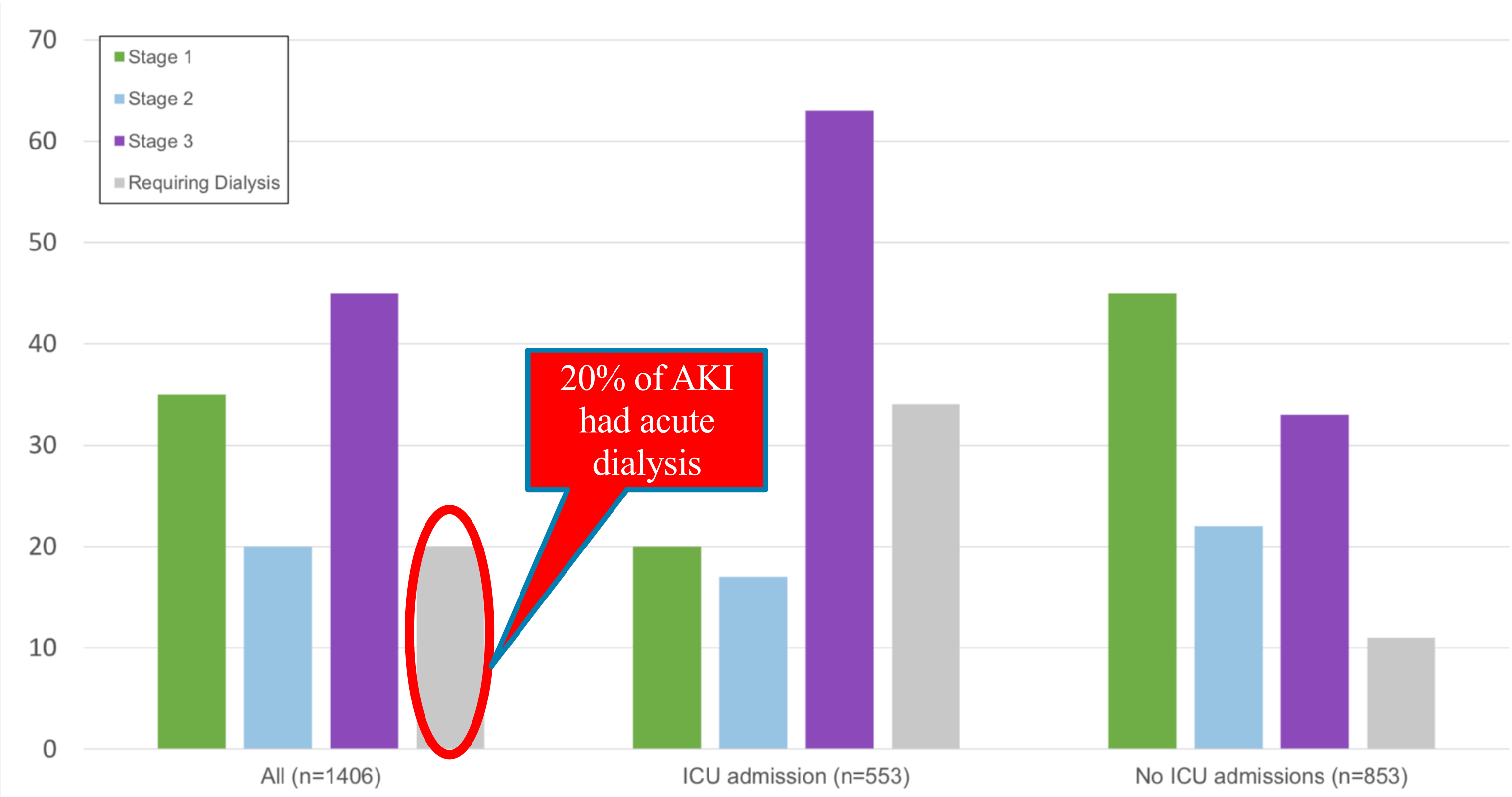
Sulaiman Somani, BS





SARS-CoV-2 is Devastating to Numerous Organ Systems

AKI Stages: Overall and ICU Admissions



Objective: To Evaluate Approaches for predicting the Need for Acute Hemodialysis over a variety of time horizons using data from <24 hours of admission

Study Workflow

A. Facilities

MSH

Training + Internal Validation

MSM MSB MSQ MSW

External Validation

B. Data

COVID - 19 Patients
(Demographics, Comorbidities, Vitals, Labs)

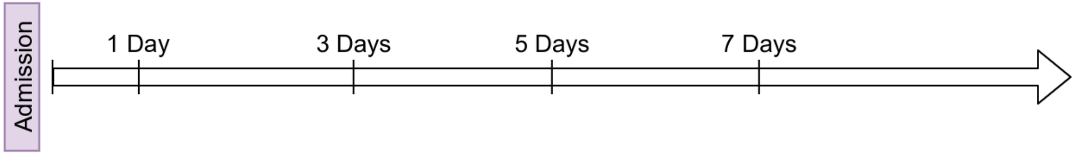
- Inpatient
- Age >= 18
- Positive COVID - 19 RT-PCR

Inclusion Criteria

- Death prior to one session of dialysis
- End Stage Renal Disease
- Features / Patients with >30% missingness
- Admitted for >48 hours before positive COVID test

Exclusion Criteria

C. Timeline



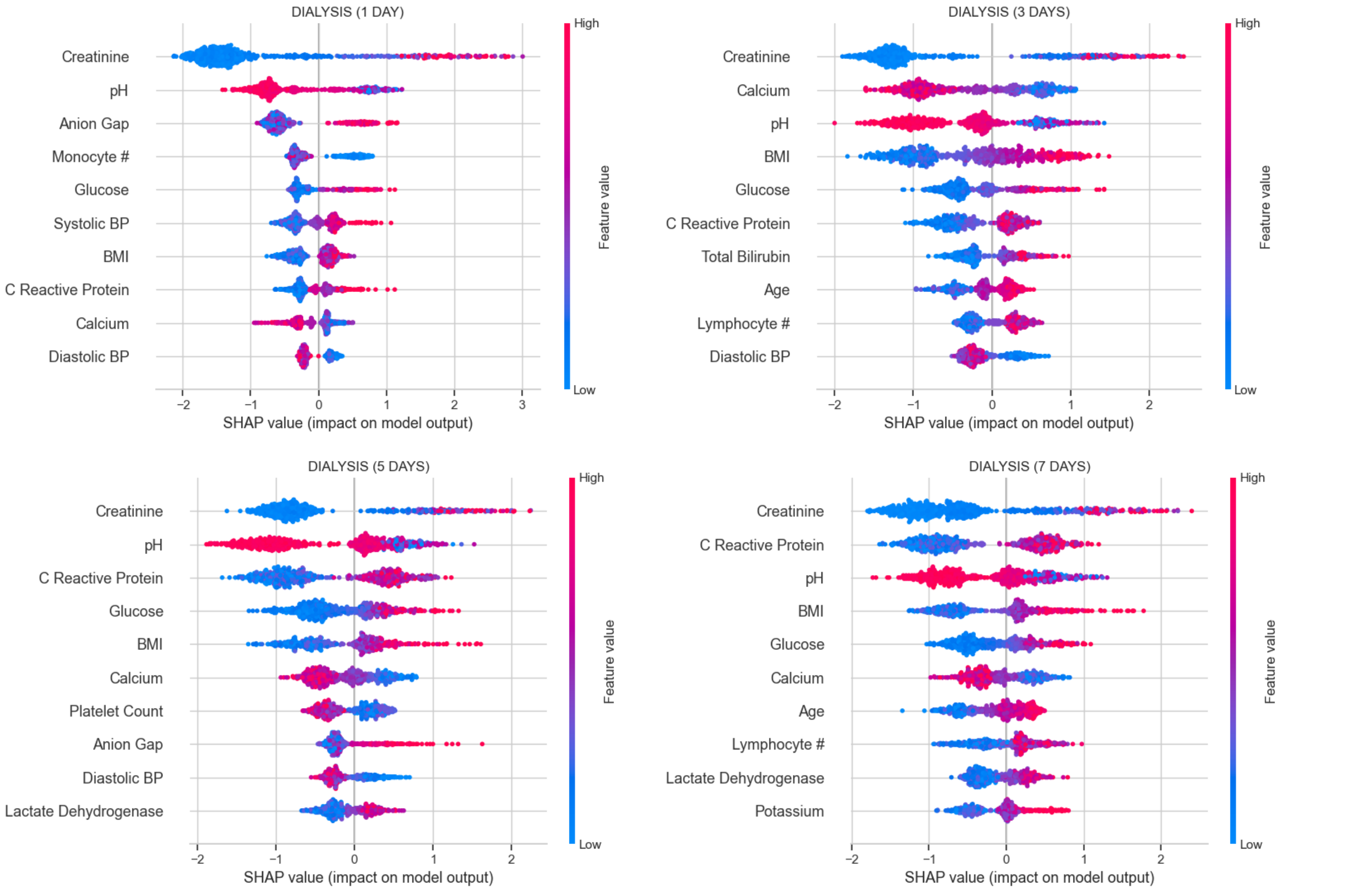
Performance Characteristics of Models over Time Horizons

	Internal Validation		External Validation	
	AUROC	AUPRC	AUROC	AUPRC
Horizon: 1 day				
LASSO	0.85	0.24	0.82	0.17
Logistic Regression	0.88	0.29	0.81	0.13
Random Forest	0.91	0.30	0.89	0.23
XGBoost (imputed)	0.93	0.34	0.91	0.30
XGBoost (not-imputed)	0.96	0.55	0.96	0.37
Horizon: 3 days				
LASSO	0.86	0.28	0.84	0.25
Logistic Regression	0.86	0.30	0.82	0.19
Random Forest	0.89	0.39	0.83	0.26
XGBoost (imputed)	0.92	0.42	0.87	0.33
XGBoost (not-imputed)	0.94	0.57	0.89	0.44

Performance Characteristics of Models over Time Horizons

	Internal Validation		External Validation	
	AUROC (95% CI)	AUPRC (95% CI)	AUROC	AUPRC
Horizon: 5 days				
LASSO	0.86	0.38	0.83	0.26
Logistic Regression	0.86	0.33	0.81	0.21
Random Forest	0.87	0.40	0.80	0.26
XGBoost (imputed)	0.87	0.43	0.86	0.32
XGBoost (not-imputed)	0.89	0.52	0.89	0.46
Horizon: 3 days				
LASSO	0.84	0.39	0.84	0.27
Logistic Regression	0.84	0.35	0.81	0.22
Random Forest	0.85	0.37	0.81	0.25
XGBoost (imputed)	0.85	0.40	0.87	0.31
XGBoost (not-imputed)	0.89	0.54	0.89	0.43

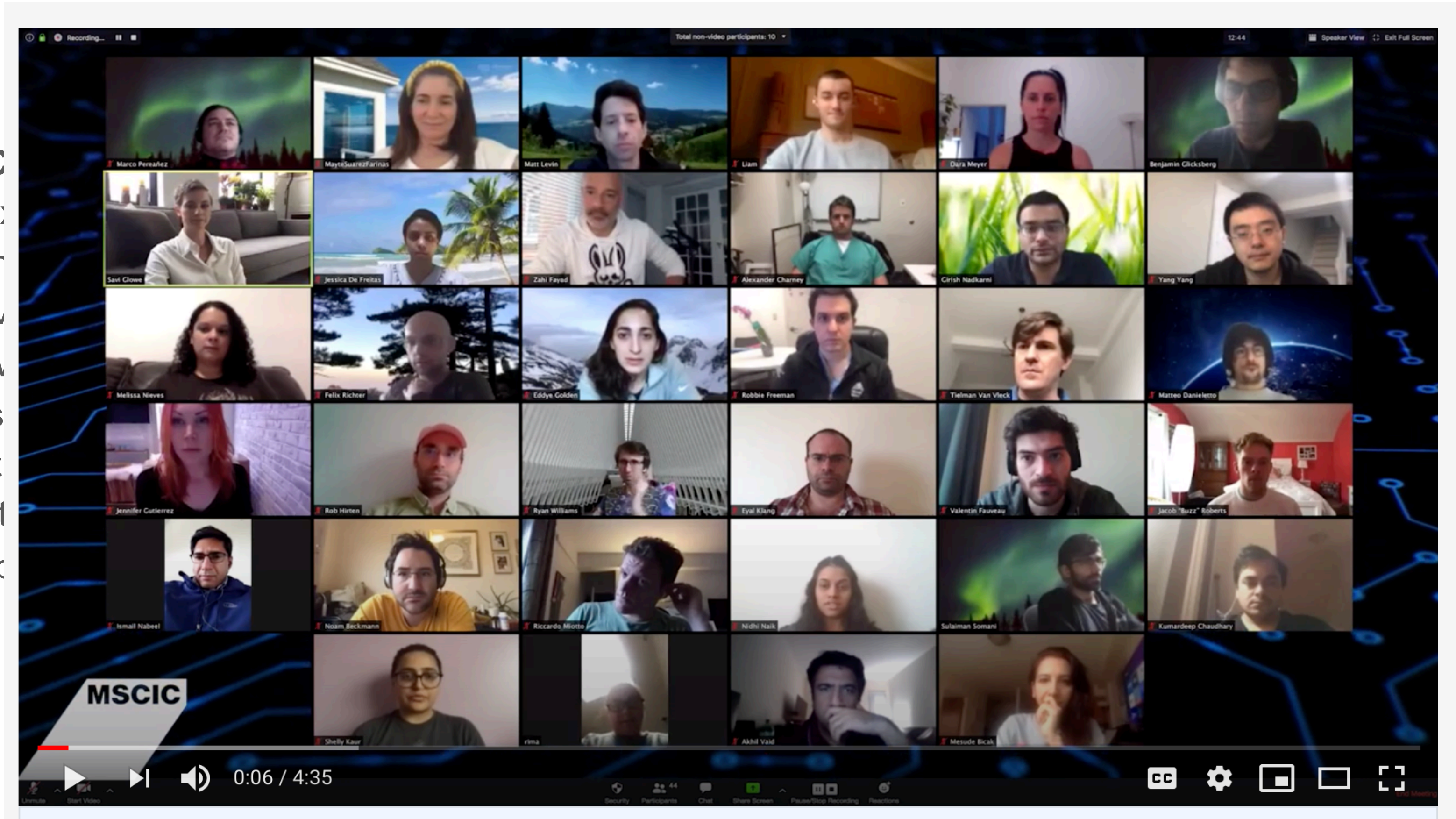
Model Explainability and Features



Acknowledgements

MSCIC

- Ale
- Zah
- Sav
- Erw
- Jos
- Pat
- Mat
- Rob



Girish Nadkarni

Paul O'Reilly