



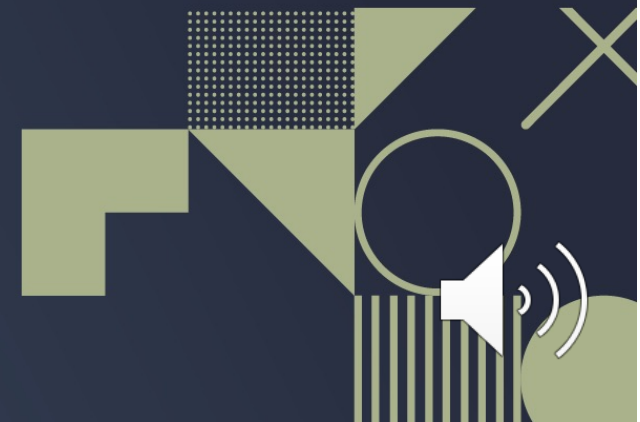
HPC Response to the COVID-19 Pandemic

Patricia Kovatch

Dean for Scientific Computing and Data

Professor, Genetics and Genomic Sciences and Systems Pharmacology

Icahn School of Medicine at Mount Sinai



About me

- Founded Scientific Computing and Data at Mount Sinai in October 2011
- Oversee HPC, clinical data warehouse, research data services
 - Director for the NIEHS-funded Human Health Exposure Analysis Resource Data Repository and Management Core
 - Co-Director for the NIA-funded Alzheimer's Disease Research Center Data Management and Statistical Core
 - Continuously funded by NIH since 2014
- Enjoying over 30 years in High Performance Computing (HPC)
 - Deployed world's 1st academic petaflop and world's 3rd fastest machine in 2009
 - Directed over 15 top 10 supercomputers for the National Science Foundation
 - Directed National Institute for Computational Sciences for the University of Tennessee at Oak Ridge National Laboratory





The Mount Sinai Health System

- School of Medicine (1968) and eight hospitals (1820)
 - Both ranked in the top 20 by US News & World Report*
- 6th largest employer in NYC (42,000 employees)
- >\$8B/year revenue
 - Over 8 million patients, 7,480 physicians, 2,000 residents/fellows
 - 3,815 hospital beds and 3.5 million outpatient visits/year
 - Over 18,000 babies born each year
 - 147 operating rooms plus 47 additional specialty ORs



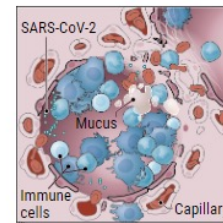
Scientific Computing and Data Mission

To enable and accelerate biomedical discovery through collaborative research using high performance computing, biomedical informatics and data science



CORONAVIRUS (COVID-19)

SARS-CoV-2 is Devastating to Numerous Organ Systems



1 Lungs

A cross section shows immune cells crowding an inflamed alveolus, whose walls break down during attack by the virus, diminishing oxygen uptake. Patients cough, fevers rise, and it takes more and more effort to breathe.

2 Liver

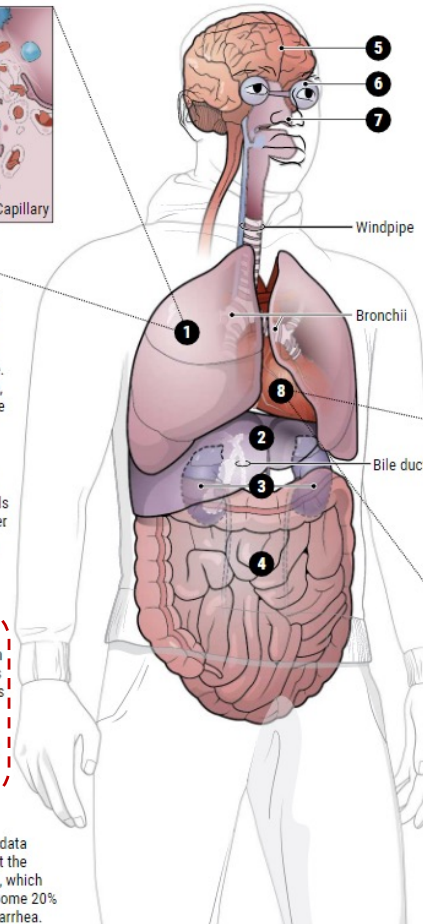
Up to half of hospitalized patients have enzyme levels that signal a struggling liver. An immune system in overdrive and drugs given to fight the virus may be causing the damage.

3 Kidneys

Kidney damage is common in severe cases and makes death more likely. The virus may attack the kidneys directly, or kidney failure may be part of whole-body events like plummeting blood pressure.

4 Intestines

Patient reports and biopsy data suggest the virus can infect the lower gastrointestinal tract, which is rich in ACE2 receptors. Some 20% or more of patients have diarrhea.



5 Brain

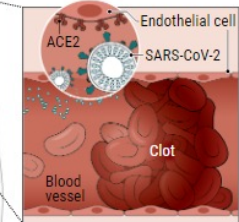
Some COVID-19 patients have strokes, seizures, mental confusion, and brain inflammation. Doctors are trying to understand which are directly caused by the virus.

6 Eyes

Conjunctivitis, inflammation of the membrane that lines the front of the eye and inner eyelid, is more common in the sickest patients.

7 Nose

Some patients lose their sense of smell. Scientists speculate that the virus may move up the nose's nerve endings and damage cells.

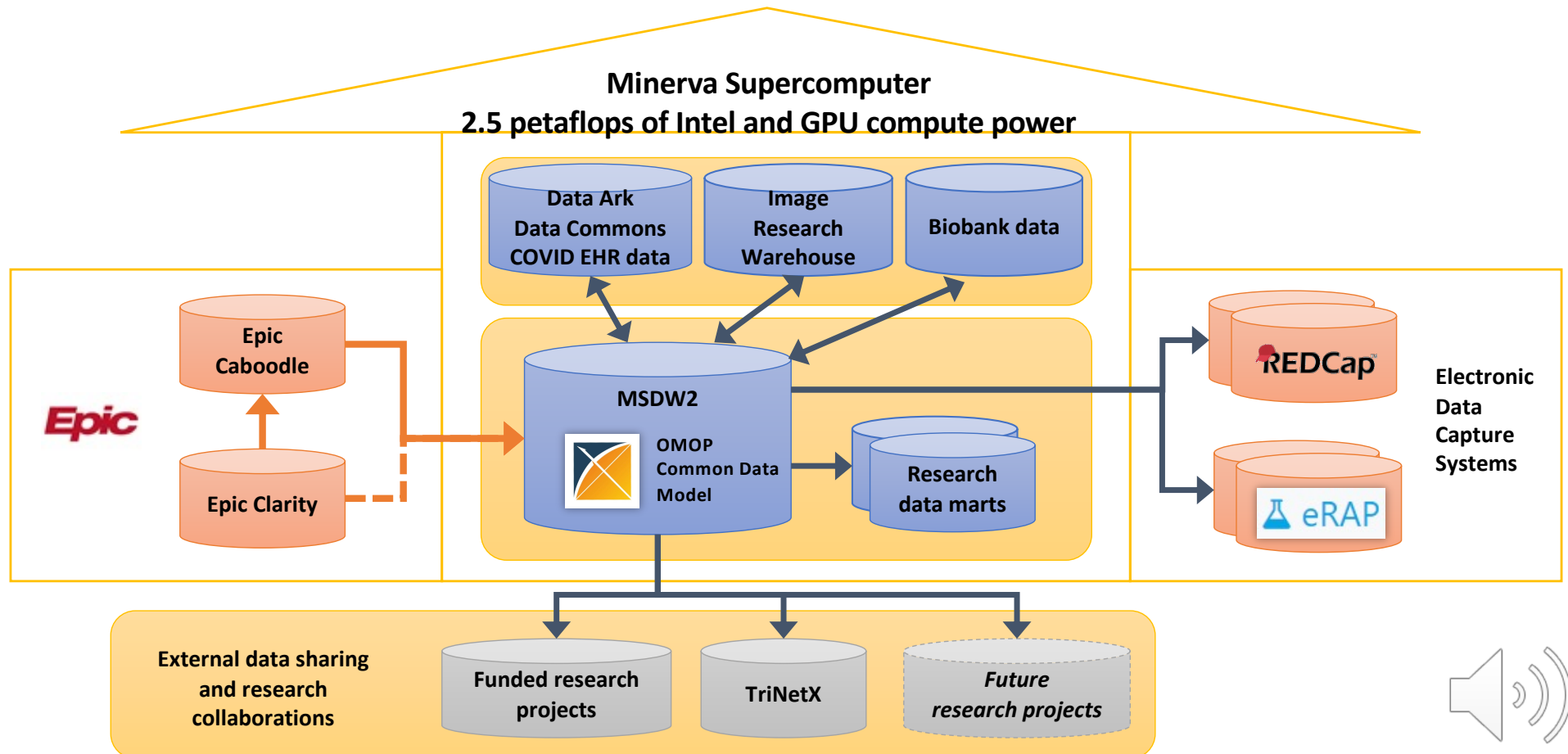


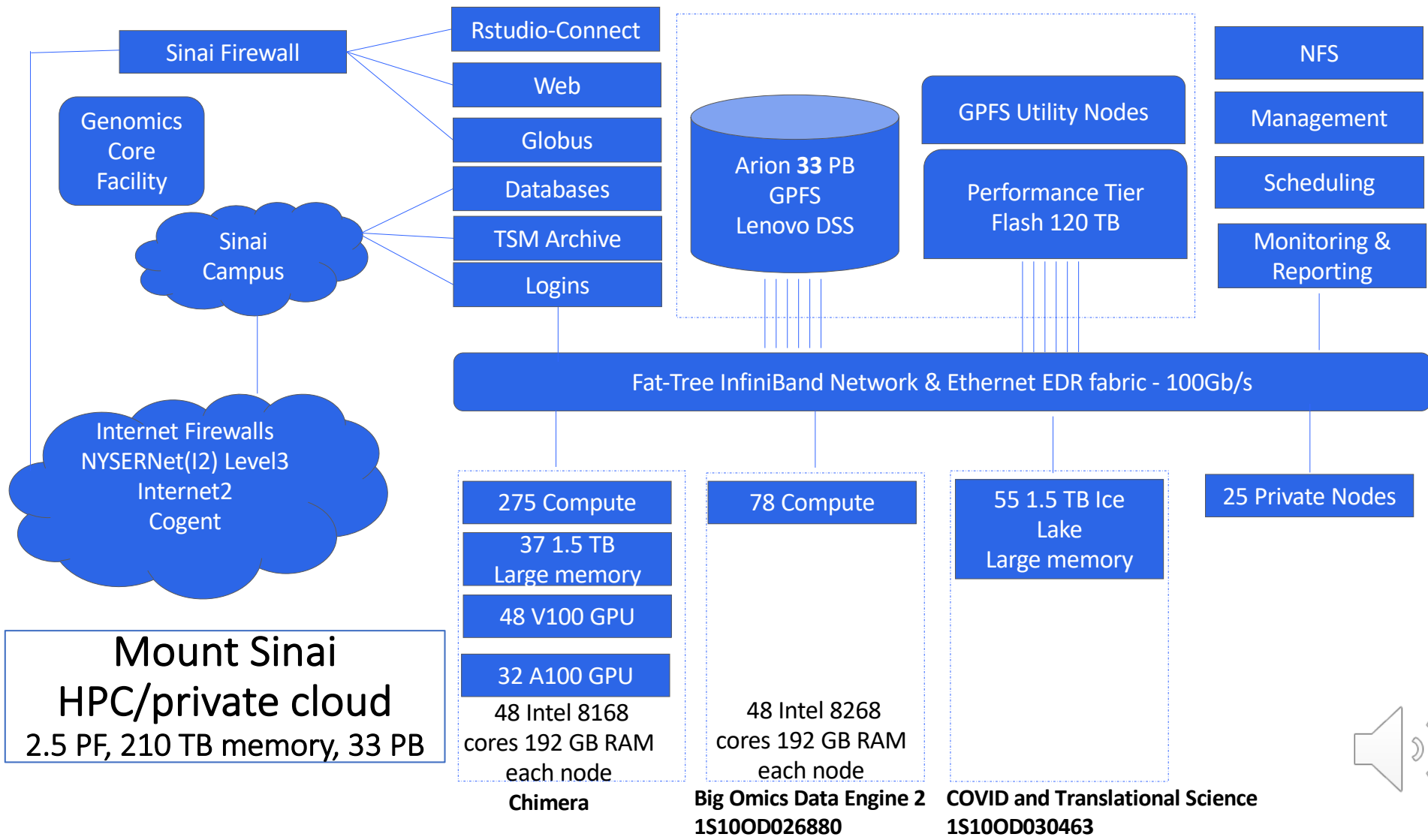
8 Heart and blood vessels

The virus (green) enters cells, likely including those lining blood vessels, by binding to ACE2 receptors on the cell surface. Infection can also promote blood clots, heart attacks, and cardiac inflammation.

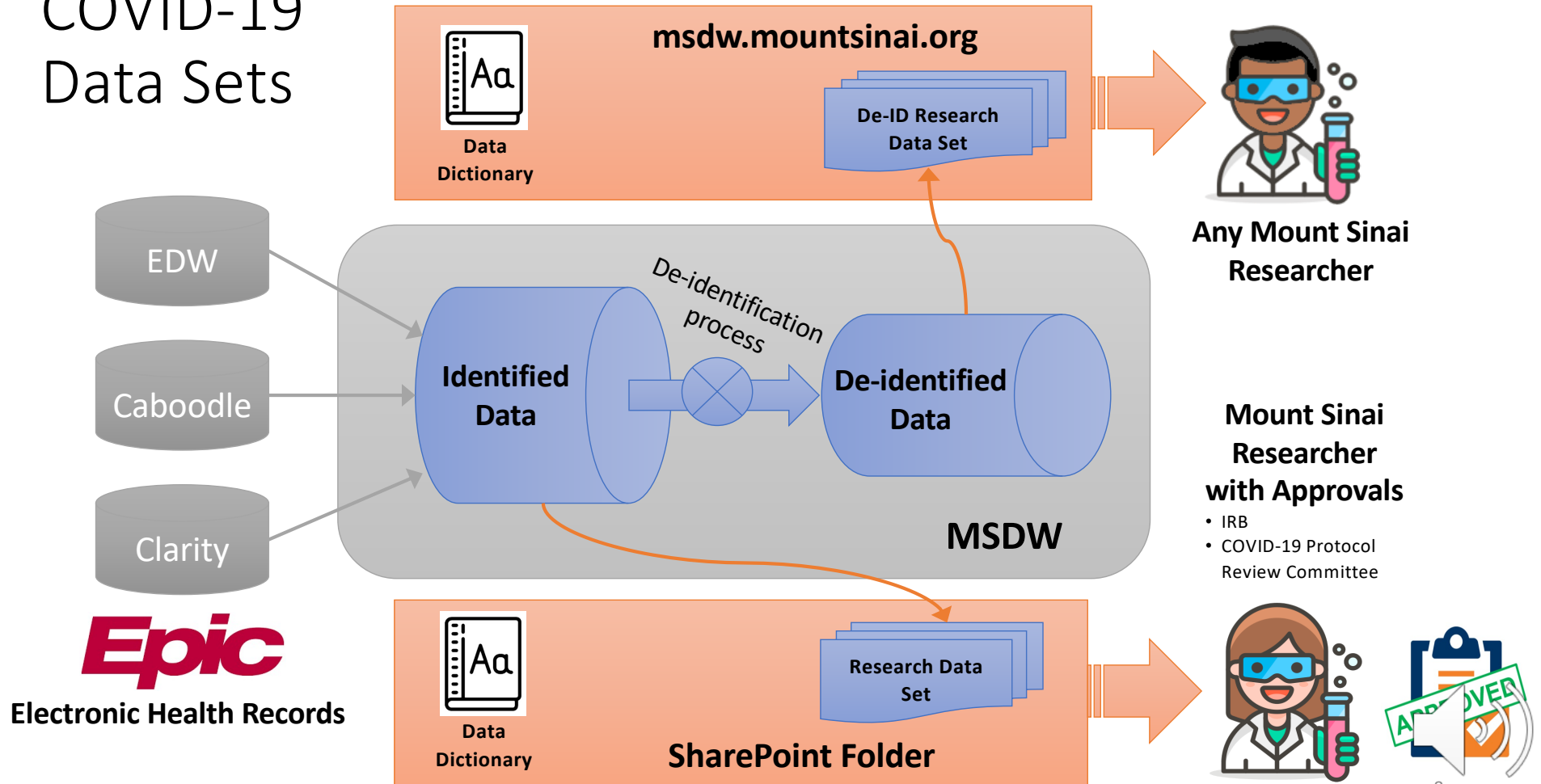


Integrated Computational and Data Ecosystem





COVID-19 Data Sets



Data de-identification for US Health Insurance Portability And Accountability (HIPAA) Privacy Rule Compliance

► **Process for COVID-19 Data Sets**

- Mask or remove all 18 Personal Health Identifiers (PHI)
- Birth date converted to age at encounter, with all ages over 89 set to 90
- Convert all dates to elapsed days relative to encounter date
- Zip codes truncated to 3 digits (or masked if population < 20k)

► **Process for Other medical record data sets**

- Same as above, but all dates for each patient are shifted by a random integer to preserve chronology

1. Patient names
2. Geographic elements like address street, city, county, zip code (first 3 digits allowed)
3. All elements of any dates (except year), including dates of birth, admission, discharge, death, and exact ages over 89
4. Telephone numbers
5. Fax numbers
6. E-mail addresses
7. Social Security Number (SSN)
8. Medical Record Number (MRN)
9. Health insurance beneficiary numbers
10. Account numbers
11. Certificate/license numbers
12. Vehicle identifiers, including license plate numbers
13. Device identification and serial numbers
14. Digital identifiers, including Uniform Resource Locators (URLs)
15. Internet Protocol (IP) addresses
16. Biometric identifiers, including finger, retinal, and voiceprints
17. Full-face photographic images and other comparable images
18. Any other uniquely identifying numbers or codes



Scientific Computing De-identified COVID-19 Research Data Set

Over 7,000 downloads
at Mount Sinai since April 2020

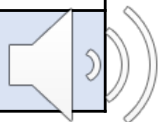
Data Sources: Epic, NYSDOH Wadsworth Lab
Inclusion Criteria: Visit Type OR Encounter Dx OR Lab Test Order/Result related to COVID-19
Purpose: Research and clinical questions

Patient Encounters <ul style="list-style-type: none"> Demographics Age, Sex, Race, Ethnicity, Zip Code (3 digits), Preferred Language Outcomes Infection Status, Mortality 		Comorbidities/Outcomes Asthma, COPD, CKD, HTN, Obesity, DM, HIV, Smoking, Cancer, CAD, CHF, OSA, A Fib, Liver Disease, AKI, VTE, Cerebral Infarction		Encounter Encounter Type, Visit Type, Patient Class, Department, Location of Care, Admission, Discharge	
Vital Signs <ul style="list-style-type: none"> BMI Max Temp Heart Rate Respiratory Rate Max Blood Pressure Min Blood Pressure Min O2 Saturation 	Lab Tests <ul style="list-style-type: none"> SARS-CoV-2 PCR, Antibody Assay IL-6, IL-8, IL-1 Beta TNF Alpha D-Dimer Ferritin LDH Fibrinogen C-Reactive Protein Procalcitonin WBC, RBC Blood Culture AST ALT Alk Phos Sodium 	<ul style="list-style-type: none"> Potassium eGFR Hemoglobin BUN Creatinine Lymphocyte Eosinophil Basophil Monocyte Neutrophil Platelet Hematocrit Troponin aCL Antibody B2GPI Antibody Phospholipid Antibody ESR 	<ul style="list-style-type: none"> BNP MCV, MCH, MCHC MPV PTT, PT, INR Albumin Calcium Uric Acid Bilirubin pH pCO2, pO2 O2 Sat HCO3 CK, CK-MB Troponin I 	Medication Administrations <ul style="list-style-type: none"> Tocilizumab Remdesivir Sarilumab Hydroxychloroquine Anakinra Azithromycin Rivaroxiban Apixaban Enoxaparin tPA Heparin 	<ul style="list-style-type: none"> Etanercept Nitric oxide Dopamine Vasopressin Norepinephrine Epinephrine Milrinone Dobutamine Phenylephrine Methylprednisolone Prednisone Dexamethasone



Select COVID-19 Data Set Research Projects

Genetics	Social and genetic determinants for COVID-19 disease susceptibility	Noura Abul-Husn, MD, PhD Eimear Kenny, PhD
	COVID-19 Predictive Modeling & Patient Tool	Girish Nadkarni, MD Benjamin Glicksberg, PhD
	Modeling of COVID Course of Care	Matthew Levin, MD
Emergency Medicine	National NHLBI ED COVID registry with PETAL/ CORAL	Lynne Richardson, MD
Cardiology	Cardiac arrhythmias in patients with COVID-19	Vivek Reddy, MD
Population Health	CDRN COVID-19 database for INSIGHT	Carol Horowitz, MD, MPH
Internal Medicine	Determining COVID-19 Natural History and Predicting Adverse Clinical Outcomes	Juan Wisnivesky, MD DrPH Keith Sigel, MD
Ob/Gyn	COVID perinatal quality database	Elizabeth Howell, MD



COVID360 Real-time Dashboard

Goal

To combine electronic medical record information, risk factors, and the output of a predictive model into an easy-to-use tool for rapid insights

The screenshot shows the COVID360 Real-time Dashboard within the Epic EMR system. The dashboard is organized into several sections:

- Patient Information:** Displays patient details such as Sex (F), DOB (77 yrs), MRN, and Visit ID.
- Risk Factors:** Lists risk factors including BMI > 40, Asthma, and COPD.
- COVID-19 Specific Health Information:** Contains a section for Risk Prediction Scores with four key metrics:
 - 72 Hour Discharge:** Likelihood of discharge is **LOW** (As of: 01-25-2021 07:35:31).
 - 24 Hour Intubation:** Likelihood of intubation is **Patient Intubated** (No Prediction).
 - 24 Hour ICU Admission:** Likelihood of ICU Admission is **Patient in ICU** (No Prediction).
 - 120 Hour Dialysis:** Likelihood of dialysis is **HIGH** (As of: 01-25-2021 07:35:31).
- Sars CoV2 PCR Test Results:** A table showing test results and dates:

Result	Result Date
DETECTED	01-19-2021 07:24:00
NOT DETECTED	12-28-2020 08:52:00
NOT DETECTED	12-12-2020 12:40:00
NOT DETECTED	11-20-2020 07:39:00
- Respiratory Insights:** Displays ventilator settings: Device Name: Ventilator, FIO2 Set: 39.0, O2 Sat: 93.0, Device Time: 01-25-2021 09:00:00.

At the bottom of the dashboard, there is a **VtSet** value of 18.0. A speaker icon is visible in the bottom right corner of the dashboard area.

Select COVID-19 Research Projects using HPC

Tracking & Surveillance	Nextstrain COVID-19 tracking	Harm Van Bakel
	Computational support for HPI COVID-19 initiative	Girish Nadkarni
Drugs & Therapeutics	Systems to identify and validate novel therapeutics for COVID-19	Bin Zhang
	Drug repurposing and crowdsourcing for COVID-19	Avi Ma'ayan
Patient Outcomes	Provide a web application to analyze & predict COVID-19 patient outcomes	Towfique Raj
	Fast-tracking COVID-19 detection test by fluidigm microfluidics technology	Stuart Sealfon
	Gene Burden Analysis of BioMe™ Donor Whole Exome Sequencing Data	Yuval Itan
Molecular Dynamics	COVID-19 viral replication and host immune response	Ming-Ming Zhou, Martin Walsh, Marta Filizola, Adolfo Garcia-Sastre
	Small Molecule Inhibitors Blocking Virus-ACE2 Interactions	Marta Filizola, Jian Jin, Florian Krammer, Adolfo Garcia-Sastre



Coronavirus in N.Y.: Manhattan Woman Is First Confirmed Case in State

A woman in her 30s who had traveled in Iran contracted the virus and is now isolated at home.

The New York Times



Early SARS-CoV-2 transmission in New York City linked to Europe

We analyzed the viral genomes of these cases and compared them to the background sequences deposited by other countries across the world in a massive global sequencing effort.

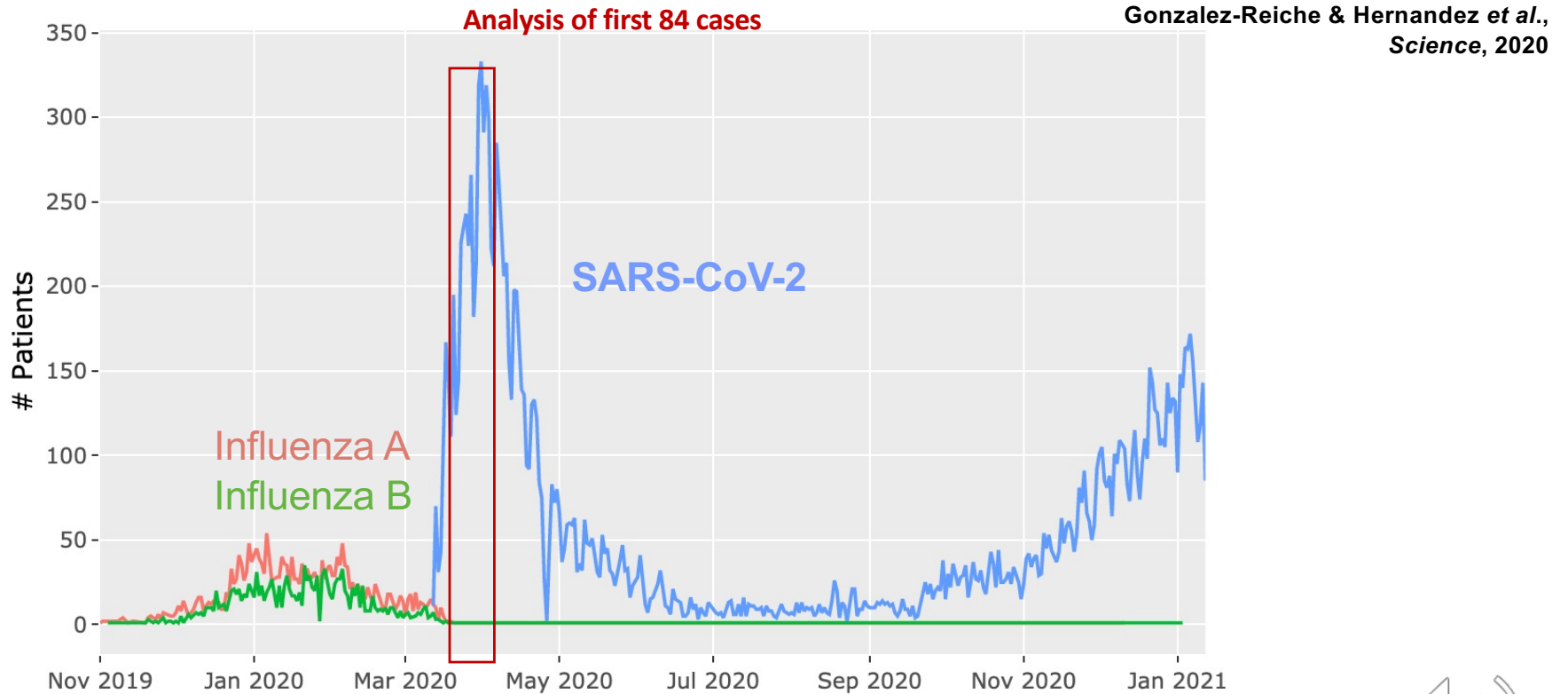
The picture that emerged from there is that the initial introduction of the virus into NY mainly came through Europe where there had been a large peak during a time when travel restrictions were not in place yet.

Gonzalez-Reiche & Hernandez et al.
Science, 2020



Genomic surveillance of SARS-CoV-2 across the Mount Sinai Health System

Daily positive SARS-CoV-2 tests



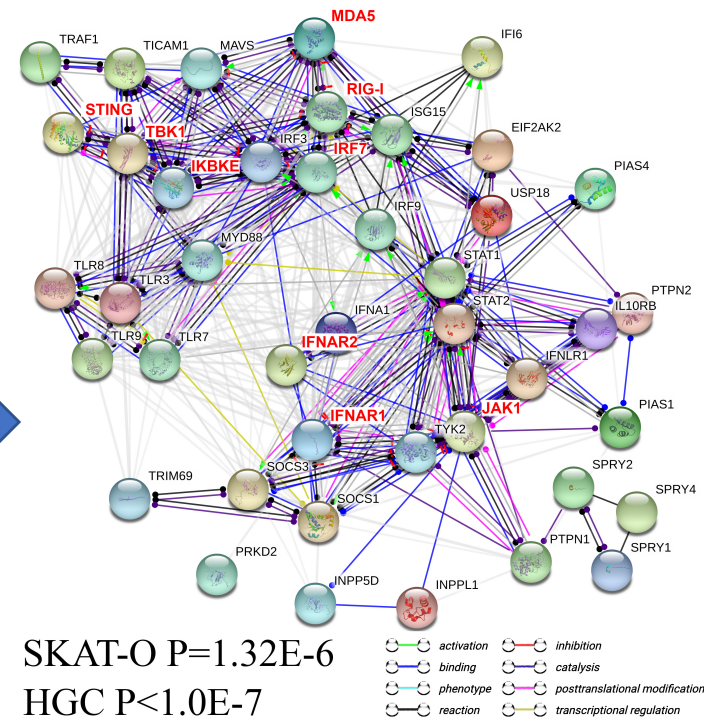
This shows a plot of the positive COVID cases along with a typical flu season to give a perception of the scale of the pandemic; within weeks we have as many cases as normally seen in an entire flu season.



COVID-19 gene burden analyses in BioME biobank

- Severe (hospitalized+ventilated): 252
- Mild (non-hospitalized): 464
- Negative: 1,268
- Now >2,500 COVID-19-positive in BioMe

p-SKAT(COVID_all vs BioMe)HM_impact	p-SKAT(COVID_hosp vs 1000G)High_impact
p-Fisher's(COVID_all vs BioMe)HM_impact	p-Fisher's(COVID_hosp vs 1000G)High_impact
p-SKAT(COVID_all vs BioMe)High_impact	p-SKAT(COVID<50 vs 1000G)HM_impact
p-Fisher's(COVID_all vs BioMe)High_impact	p-Fisher's(COVID<50 vs 1000G)HM_impact
p-SKAT(COVID_hosp vs BioMe)HM_impact	p-SKAT(COVID<50 vs 1000G)High_impact
p-Fisher's(COVID_hosp vs BioMe)HM_impact	p-Fisher's(COVID<50 vs 1000G)High_impact
p-SKAT(COVID_hosp vs BioMe)High_impact	p-SKAT(UKBB_hosp vs UKBB_nonhosp)HM_impact
p-Fisher's(COVID_hosp vs BioMe)High_impact	p-Fisher's(UKBB_hosp vs UKBB_nonhosp)HM_impact
p-SKAT(BioMe_hosp vs BioMe_nonhosp)HM_impact	p-SKAT(UKBB_hosp vs UKBB_nonhosp)High_impact
p-Fisher's(BioMe_hosp vs BioMe_nonhosp)HM_impact	p-Fisher's(UKBB_hosp vs UKBB_nonhosp)High_impact
p-SKAT(BioMe_hosp vs BioMe_nonhosp)High_impact	p-SKAT(UKBB_allpos vs 1000G)HM_impact
p-Fisher's(BioMe_hosp vs BioMe_nonhosp)High_impact	p-Fisher's(UKBB_allpos vs 1000G)HM_impact
p-SKAT(COVID<50 vs BioMe)HM_impact	p-SKAT(UKBB_allpos vs 1000G)High_impact
p-Fisher's(COVID<50 vs BioMe)HM_impact	p-Fisher's(UKBB_allpos vs 1000G)High_impact
p-SKAT(COVID<50 vs BioMe)High_impact	p-SKAT(UKBB_hosp vs 1000G)HM_impact
p-Fisher's(COVID<50 vs BioMe)High_impact	p-Fisher's(UKBB_hosp vs 1000G)HM_impact
p-SKAT(COVID<50 vs nonhosp_BioMe)HM_impact	p-SKAT(UKBB_hosp vs 1000G)High_impact
p-Fisher's(COVID<50 vs nonhosp_BioMe)HM_impact	p-Fisher's(UKBB_hosp vs 1000G)High_impact
p-SKAT(COVID<50 vs nonhosp_BioMe)High_impact	p-SKAT(UKBB_hosp vs UKBB_neg)HM_impact
p-Fisher's(COVID<50 vs nonhosp_BioMe)High_impact	p-Fisher's(UKBB_hosp vs UKBB_neg)HM_impact
p-SKAT(COVID_hosp vs 1000G)HM_impact	p-SKAT(UKBB_hosp vs UKBB_neg)High_impact
p-Fisher's(COVID_hosp vs 1000G)HM_impact	p-Fisher's(UKBB_hosp vs UKBB_neg)High_impact



- Analyses performed at the Yuval Itan lab:
 - SKAT-O analyses done by Yiming Wu
 - HGC simulations done by Cigdem Sevim Bayrak

- 43 Interferon genes (IFN, with Dusan Bogunovic): $P=1.32 \times 10^{-6}$
 - $P=3.04 \times 10^{-4}$ in UKBB
- 21 ACE pathway genes: 3.26×10^{-6}
 - $P=2.91 \times 10^{-5}$ in UKBB



There is a group of genes that contain relatively more mutations in severe COVID patients when compared to controls, which were identified by a method called gene burden analysis.

Impact of COVID-19 on individuals with Down syndrome (DS)

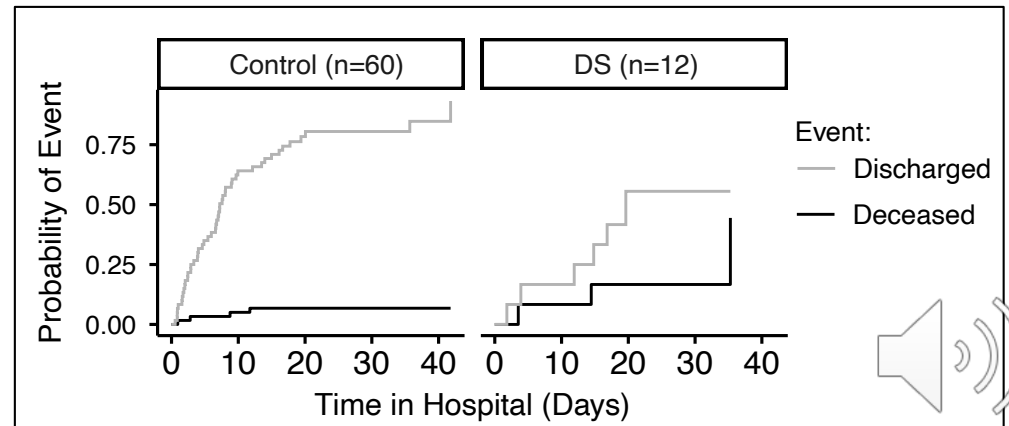
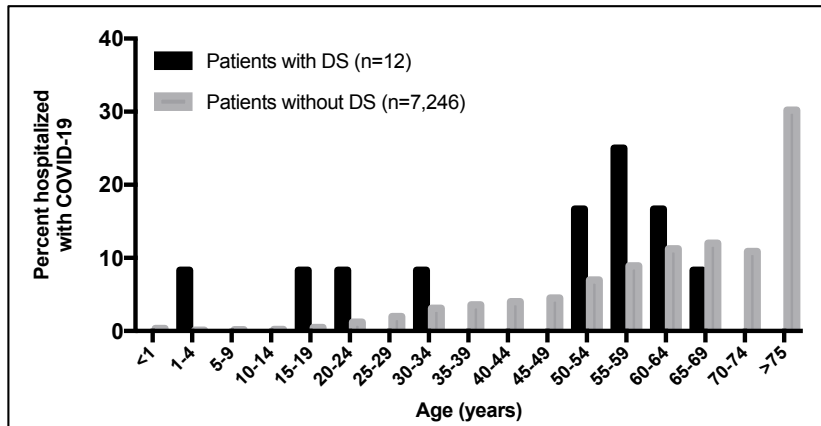
- In patients who were hospitalized with COVID-19, those with DS were on average 10 years younger than those without DS
- Patients with DS had more severe disease than age-matched controls, particularly an increased incidence of sepsis and mechanical ventilation
- An appreciably higher proportion of patients died in the DS group than in the control group

© American College of Medical Genetics and Genomics

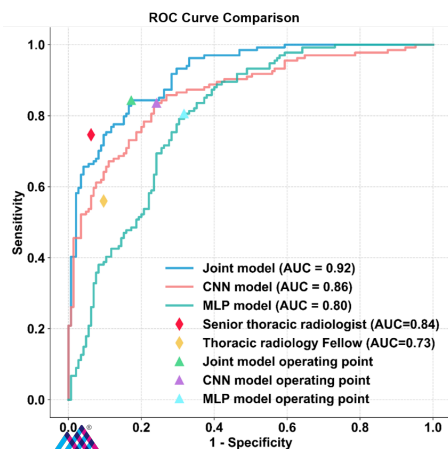
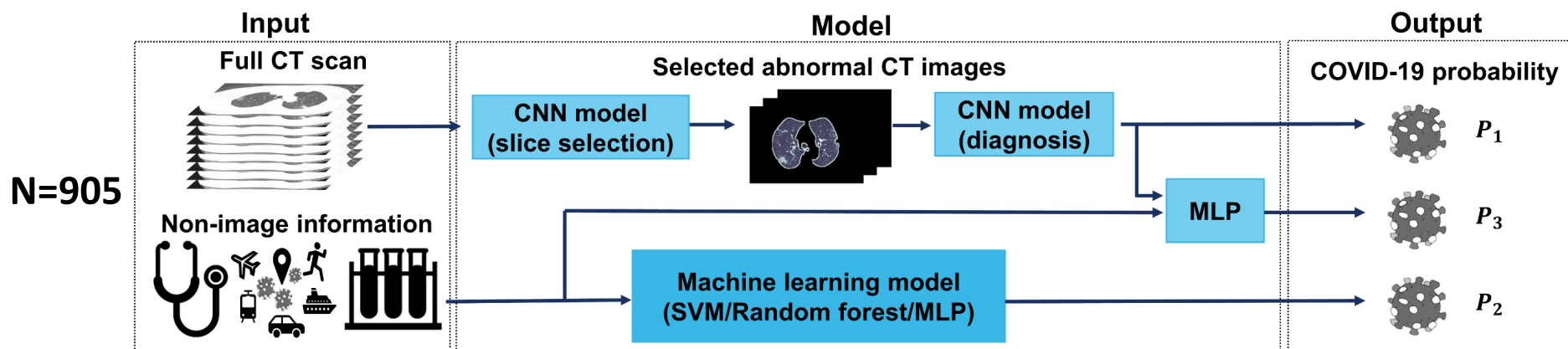
BRIEF COMMUNICATION | Genetics in Medicine

Individuals with Down syndrome hospitalized with COVID-19 have more severe disease

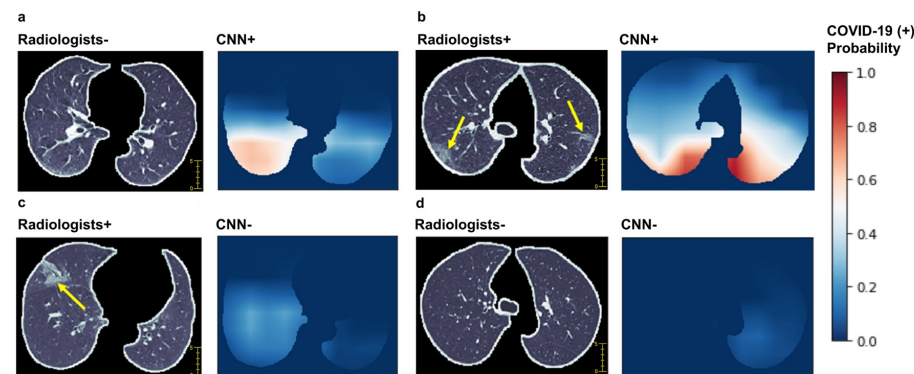
Louise Malle, BA^{1,2,3,4}, Cynthia Gao, BS⁵, Chin Hur, MD, MPH⁶, Han Q. Truong, BS⁶, Nicole M. Bouvier, MD^{1,5}, Bethany Percha, PhD⁶, Xiao-Fei Kong, MD, PhD⁶ and Dusan Bogunovic, PhD^{1,2,3,4,7}



Artificial intelligence-enabled rapid diagnosis of COVID-19 patients



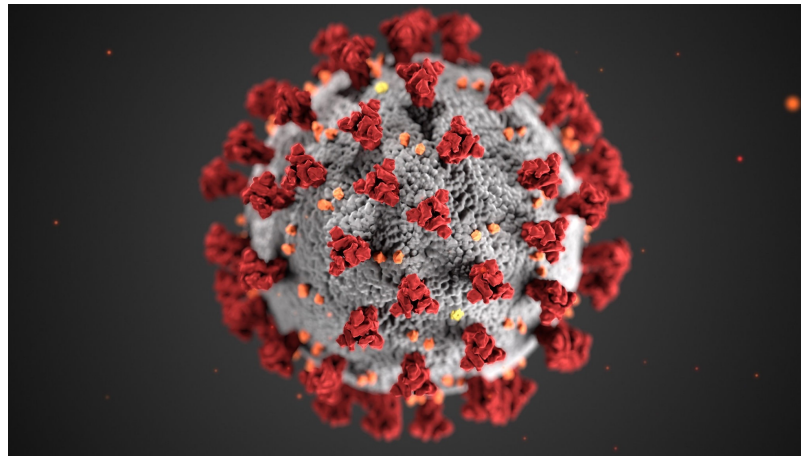
- The AI system achieved an area under the curve of 0.92 and had equal sensitivity as compared to a senior thoracic radiologist.
- The AI system also improved the detection of patients who were positive for COVID-19 via RT-PCR who presented with normal CT scans, correctly identifying 17 of 25 (68%) patients, whereas radiologists classified all of these patients as COVID-19 negative.
- When CT scans and associated clinical history are available, the proposed AI system can help to rapidly diagnose COVID-19 patients (faster than RT-PCR).



Mei, X; Yang Y; Fayad ZA et al. Nat Med



National Institutes of Health (NIH) funded R01 DA04547-02S1:
“A structure-based drug discovery effort to combat COVID-19
and future outbreaks”



Aim 1. To identify potential virus-ACE2 disruptors among FDA-approved drugs.

Aim 2. To identify novel chemotypes as backup or additional virus-ACE2 interface disruptors.

Collaboration with Marta Filizola, Adolfo Garcia-Sastre, Jian Jin, and Florian Kramer



The COVID-19 Drug and Gene Set Library

A collection of drug and gene sets related to COVID-19 research contributed by the community

355 total
355 drug sets

2003 unique drugs

575 total
575 gene sets

21818 unique genes

16 in-vitro drug screen drug sets

In-vitro drug screens to identify compounds that block SARS-CoV-2 replication in cell-based reporter assays.



106 differential expression signatures gene sets

Genome-wide differentially expressed genes in cells and tissues due to SARS-CoV-2 infection profiled with RNA-seq.



17 CRISPR screen gene sets

Genome-wide CRISPR screens to identify hits that increase or decrease SARS-CoV-2 proliferation in cell-based assays.



Drug sets | Gene sets

Experimental drug sets | Computational drug sets | Twitter drug sets | All drug sets

Draw a Venn diagram

Search in description, metadata or drugs:

Description	Drugs	DrugEnrichr link
<input type="checkbox"/> 15 SARS-CoV-2 inhibitors in Vero E6 cells from Jan et al. 2021	15 drugs	🔗
<input type="checkbox"/> 24 FDA-approved drugs inhibiting SARS-CoV-2 from Xiao et al.	24 drugs	🔗
<input type="checkbox"/> 25 compounds against SARS-CoV-2 via AlphaLisa BDD-ACE2 assay from Hanson et	25 drugs	🔗

Kuleshov MV, Stein DJ, Clarke DJB, Kropiwnicki E, Jagodnik KM, Bartal A, Evangelista JE, Hom J, Cheng M, Bailey A, Zhou A, Ferguson LB, Lachmann A, Ma'ayan A. The COVID-19 gene and drug set library. *Patterns* 1(6):100090 (2020) PMID: 32838343

<https://maayanlab.cloud/covid19/>



Thank you!

- **We have open positions!!**
 - HPC Architect
 - Computational and Data Ecosystem Manager
 - Clinical Data Specialist
 - Bioinformatician
 - Application Support Specialist
 - Healthcare SQL Analyst
 - ETL Developer

careers.mountsinai.org

