CMSS Presents:

COVID-19: Accelerating Real-Time Electronic Data Capture for Tracking, Learning and Improvement

July 8, 2020 | 2:00 – 3:30 pm ET

CMSS WEBINAR SERIES

Advancing Clinical Registries to Support Pandemic Treatment and Response

CMSS WEBINAR SERIES

Advancing Clinical Registries to Support Pandemic Treatment and Response

The series will address key questions related to the rapid development, deployment and implementation of Covid-19 focused clinical registries and clinical repositories by specialty societies and academia.

SUMMER 2020 | FREE TO ATTEND

About the Series:

- Made possible with funding from the Gordon and Betty Moore Foundation
- To foster collaboration between specialty societies and academia, we are grateful to collaborate with the Association of Academic Medical Colleges

Continue the Conversation:

- Use #COVIDRegistries when tweeting about the webinar series
- Follow @CMSSMed and visit CMSS.org for frequent updates

CMSS WEBINAR SERIES

Advancing Clinical Registries to Support Pandemic Treatment and Response

Today's Webinar:

COVID-19: Accelerating Real-Time Electronic Data Capture for Tracking, Learning and Improvement

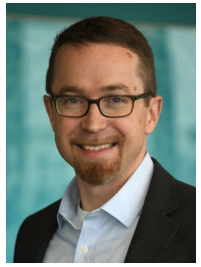
Moderator:

Atul Butte, MD, PhD
Priscilla Chan and Mark Zuckerberg Distinguished Professor; Director, Bakar Computational Health Sciences Institute, UCSF, Chief Data Scientist, UC Health

Host:

Helen Burstin, MD, MPH, MACP
Chief Executive Officer
Council of Medical Specialty Societies (CMSS)

Panelists:



Tellen D. Bennett, MD, MS
Section Head, Informatics and Data Science; Associate Professor, Department of Pediatrics, University of Colorado School of Medicine

Andrew Ip, MD, MS
Outcomes and Value Research Division, John Theurer Cancer Center; Hackensack Meridian Health

Subha Madhavan, PhD, FACMI
Chief Data Scientist, Georgetown University Medical Center

Jessie Tenenbaum, PhD
Chief Data Officer, NC Department of Health and Human Services

UNIVERSITY
OF
CALIFORNIA

Precisely Practicing Medicine from 700 Trillion Points of University of California Health Data

Atul Butte, MD, PhD

Chief Data Scientist, University of California Health (UC Health)

Director, Bakar Computational Health Sciences Institute, UCSF

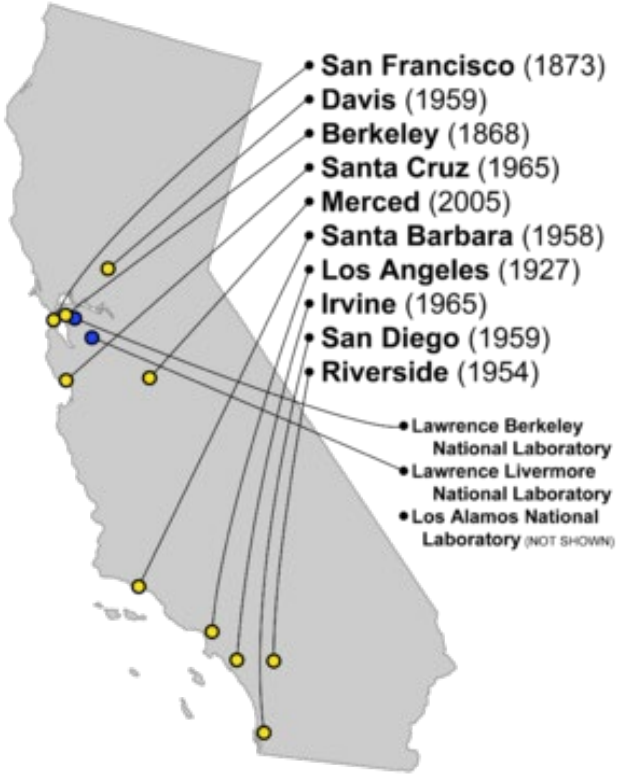
Priscilla Chan and Mark Zuckerberg Distinguished Professor

University of California

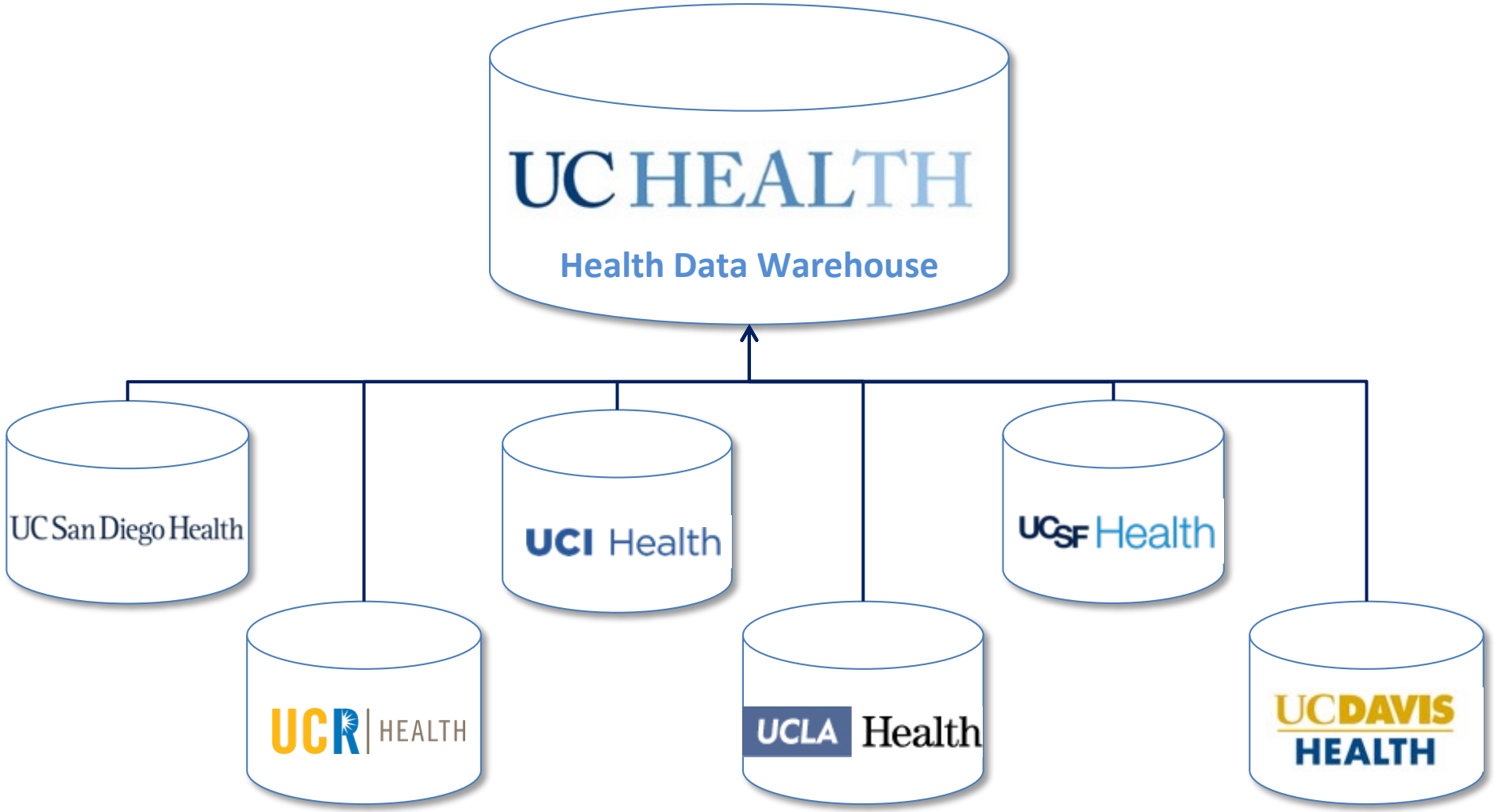
- 10 campuses and 3 national labs
- ~200,000 employees, ~250,000 students/yr

University of California Health

- 19 health professional schools (6 med schools)
- Train half the medical students and residents in California
- UCSF and UCLA are in US News top 10
- 5 NCI Comprehensive Cancer Centers, 5 NIH CTSA
- ~$2 billion NIH funding
- $13+ billion clinical operating revenue
- 5000 faculty physicians, 12000 nurses



Combining healthcare data from across the six University of California medical schools and systems

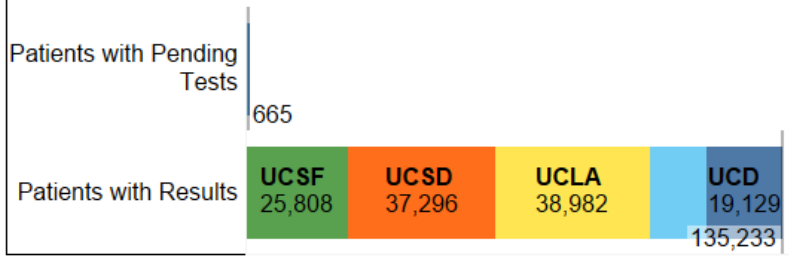


UC Health Data Analytics Platform

UC Health COVID-19 Patients (Data through 7/6/2020 11:59 PM)

Interim report, data subject to correction, UC Health patients only, tests for external partners excluded.

Cumulative Patient Test Status

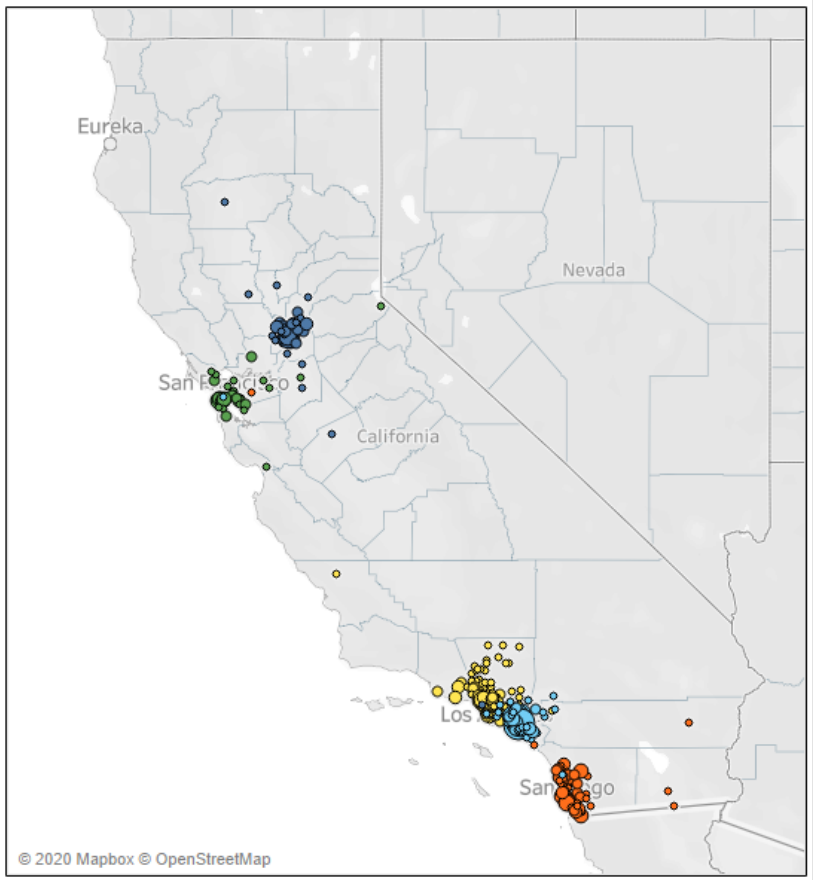


Cumulative Testing

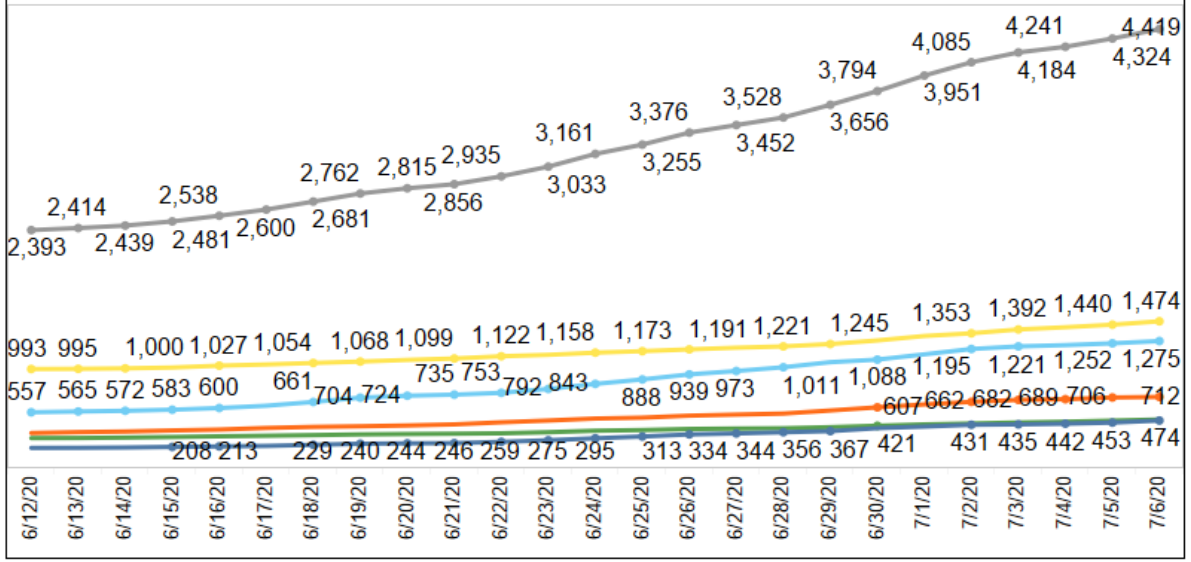
	UCD	UCI	UCLA	UCSD	UCSF	UC Health
Patients with Pending Tests	31	137	0	182	315	665
Patients with Results	19,129	14,018	38,982	37,296	25,808	135,233
Total Positives	474	1,275	1,474	712	484	4,419

Patients with Positive Test Results by Home Location

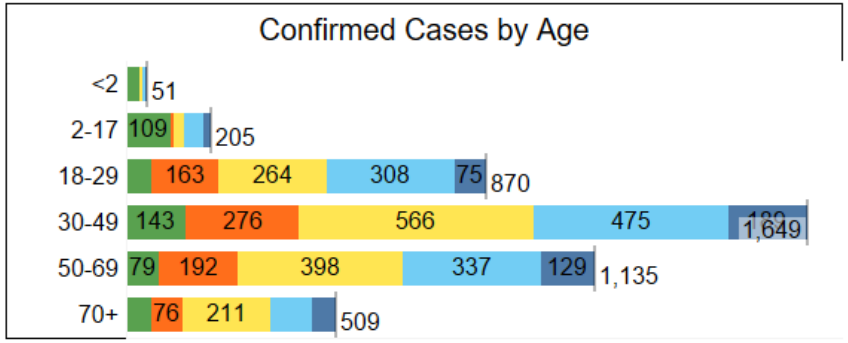
7/1/2020 12:00:00 AM to 7/7/2020 11:59:59 PM



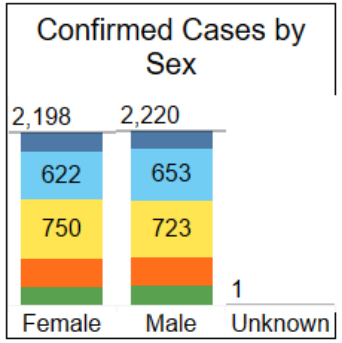
Cumulative Confirmed Cases by Location



Confirmed Cases by Age



Confirmed Cases by Sex



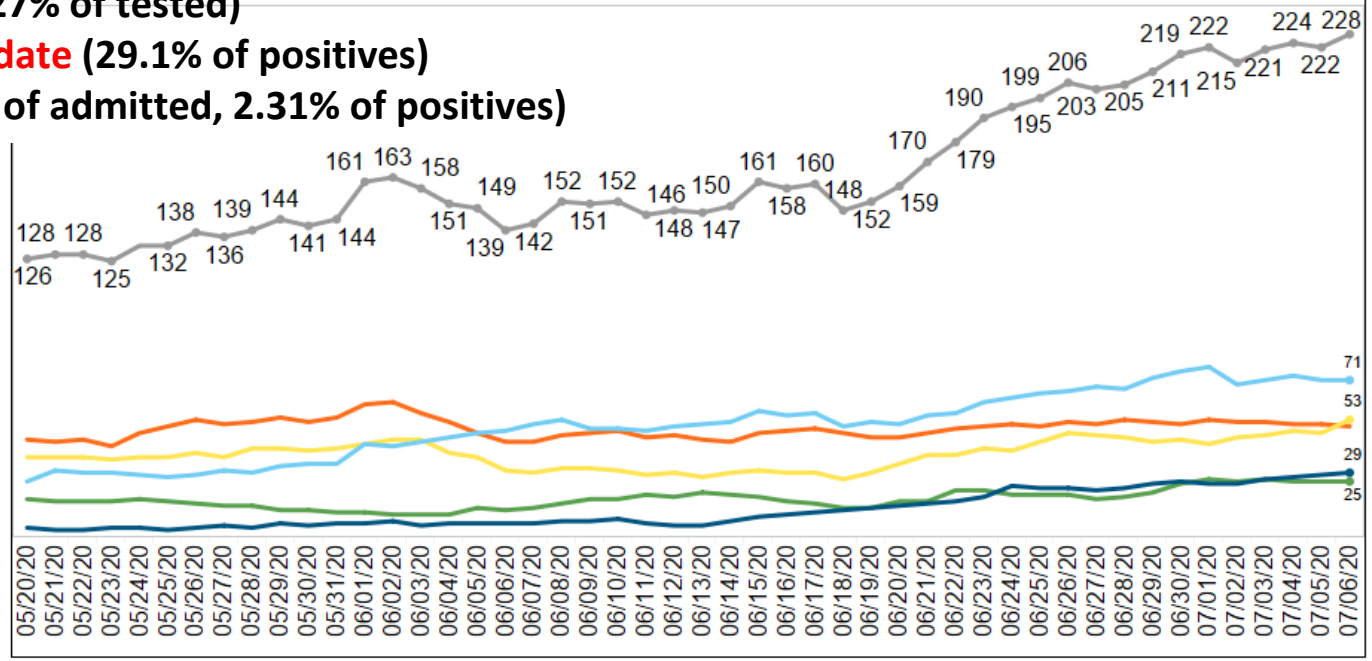
COVID-19 Inpatients (Data through 7/6/2020 11:59 PM)

Interim report, data subject to later change. Counting inpatient admitted patients, and excluding patients with pending COVID-19 tests, in emergency departments, or in observation. We are not removing patients with later resolved infections. These numbers may not directly match local dashboards.

135233 tested
4419 positives (3.27% of tested)
1285 admitted to date (29.1% of positives)
102 expired (7.9% of admitted, 2.31% of positives)

COVID-19 inpatients seen during each day

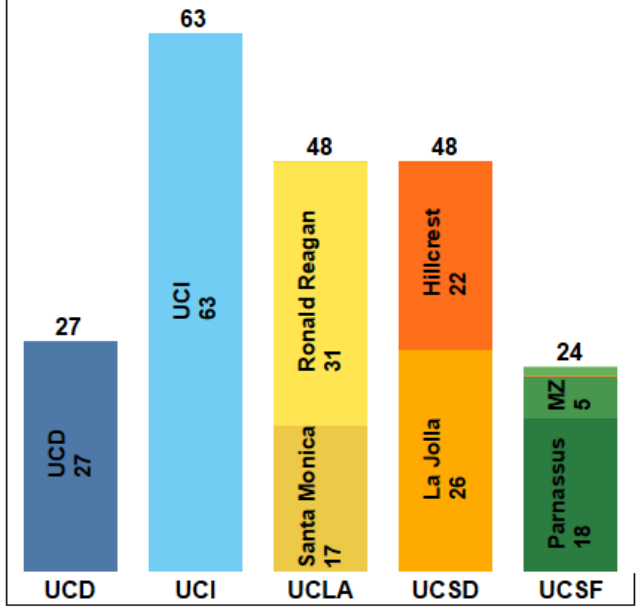
Patients are counted if present in an inpatient stay at any point during a day



COVID-19 inpatient census as of

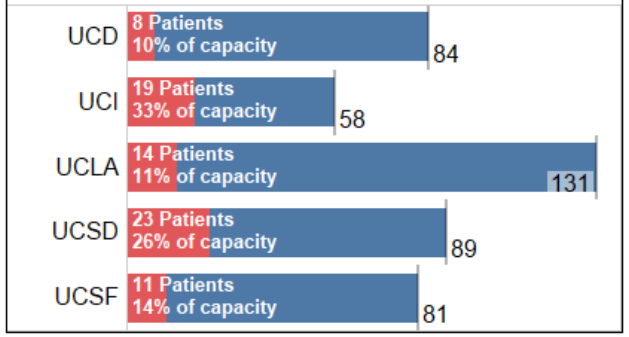
As of 11:59PM 7/6/2020

(Potentially includes inpatients with no longer detectable virus)

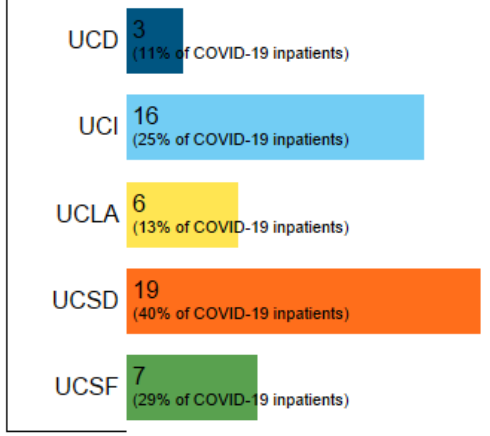


COVID-19 Patients in ICU

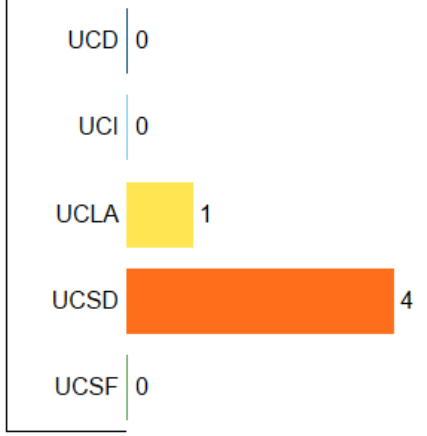
ICU status is driven by unit name and may not match with local dashboards. Total site ICU capacity indicated by blue bar and may be occupied by persons under investigation(PUI) and non-COVID-19 cases



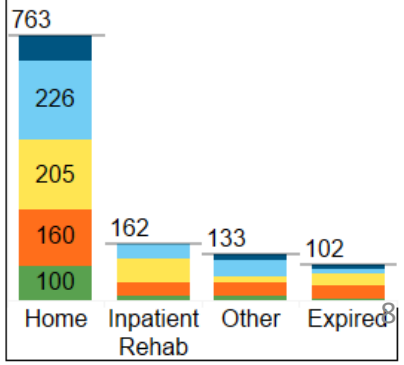
COVID-19 Patients on Ventilator



COVID-19 Patients on ECMO



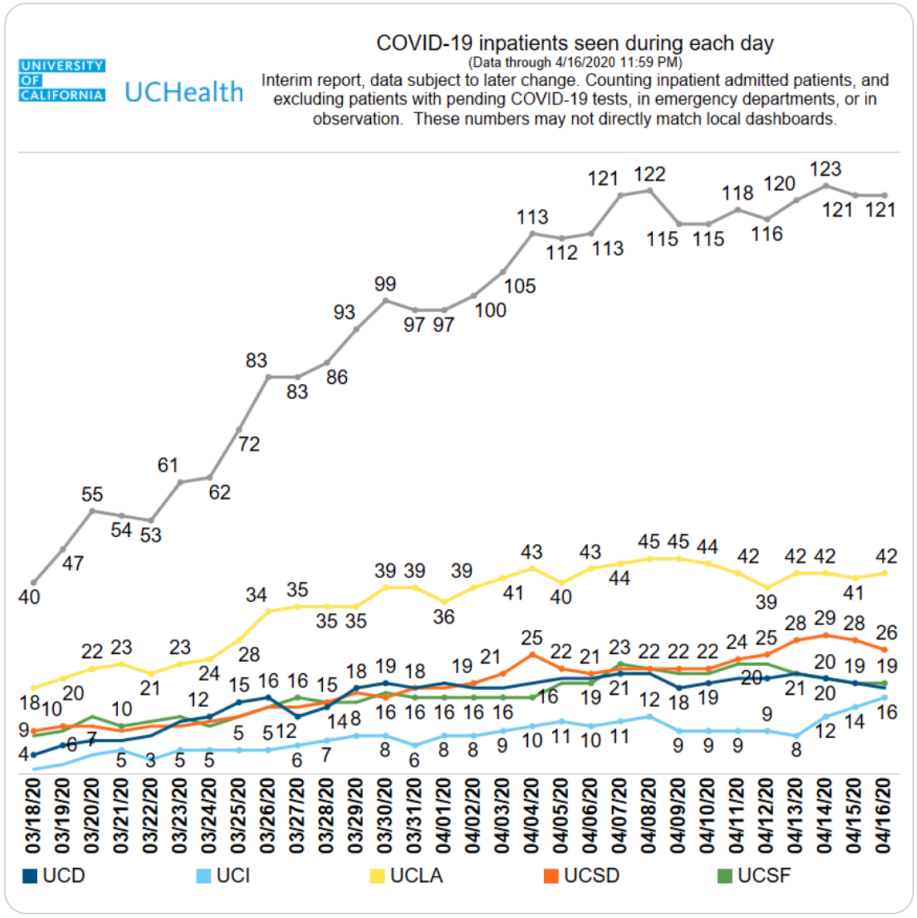
COVID-19 Inpatient Discharged Patient Dispositions



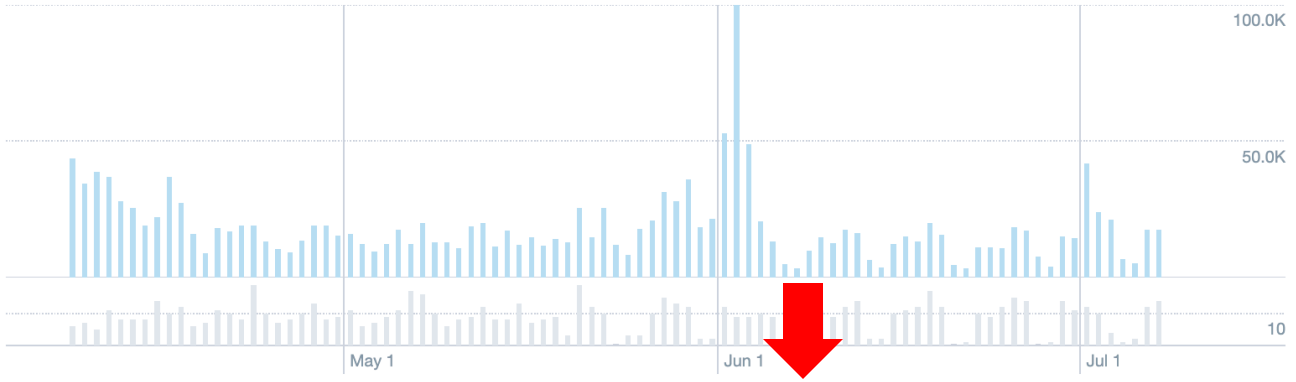
Follow @UofCAHealth for these numbers every afternoon

University of California Health
@UofCAHealth

1/6 Daily #COVID19 update: 298 #SARSCoV2 positive patients have needed admission to date in any of our 10 hospitals and 5 academic medical centers; 166 patients have been discharged home. @UofCAHealth hospitals cared for 121 #SARSCoV2 positive inpatients yesterday.



Your Tweets earned **1.8M impressions** over this **91 day** period



Tweets | Top Tweets | Tweets and replies | Promoted

	Impressions	Engagements	Engagement rate
University of California Health @UofCAHealth · Apr 11 1/5 Today's @UofCAHealth #COVID19 update: 15777 patients received #SARSCoV2 test results to date; 979 had detectable virus = 6.2%. These patients are cared for at @UCSDHealth @UCLAHealth @UCIrvineHealth @UCSFHospitals & @UCDavisHealth. pic.twitter.com/6CoFERwyQn View Tweet activity	29,984	2,669	8.9%
University of California Health @UofCAHealth · Jul 1 1/7 Daily #COVID19 update from our 5 academic med centers: 1,152 #SARSCoV2 positive patients have needed admission to date in any of our 12 hospitals; 672 patients have been discharged home. @UofCAHealth hospitals cared for 202 #SARSCoV2 positive inpatients yesterday. pic.twitter.com/O0uJ3sKOQw View Tweet activity	28,644	3,713	13.0%
University of California Health @UofCAHealth · Apr 9 1/5 Today's #COVID19 update: the 5 @UofCAHealth academic med centers (@UCSDHealth @UCLAHealth	23,819	4,323	18.1%

UC Health COVID Research Data Set (UC CORDS)

- Access open up to all UC Health research faculty, staff, students
  - Users sign a UC-wide CORDS Data Use Agreement
- Access through each campus's existing secure research environment
  - Cannot download the dataset or remove from the environment
- HIPAA Limited Data Set (deidentified, but with dates)
  - UCSF IRB has approved our UC Health Limited Data Set work as HIPAA Exempt
- All UC Health IRB directors are in agreement
  - Not Human Subjects Research (NHSR)
  - No IRB submission is required for end users
- Regenerated every Wednesday, transferred Thursday and Friday

UNIVERSITY OF CALIFORNIA Health Data Warehouse

analytics.uchealth.edu

HOME TEAM ABOUT COVID-19 **UC CORDS**

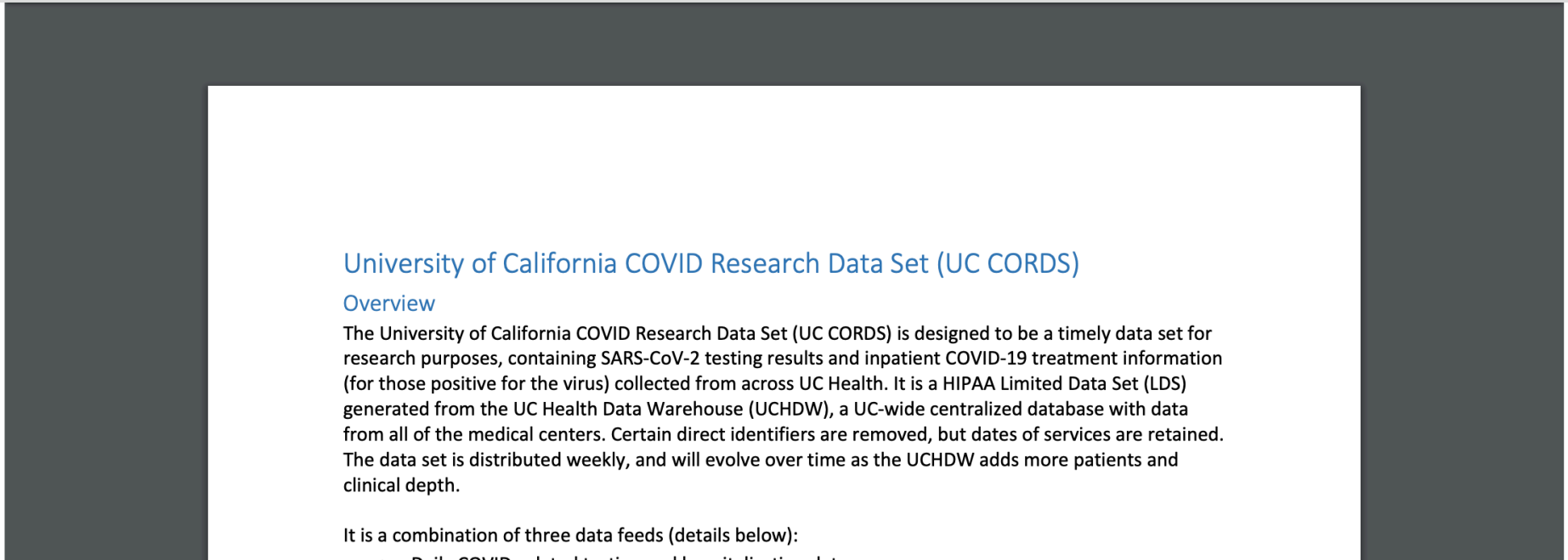
University of California COVID Research Data Set (UC CORDS)

Overview

The University of California COVID Research Data Set (UC CORDS) is designed to be a timely data set for research purposes, containing SARS-CoV-2 testing results and inpatient COVID-19 treatment information (for those positive for the virus) collected from across UC Health. It is a HIPAA Limited Data Set (LDS) generated from the UC Health Data Warehouse (UCHDW), a UC-wide centralized database with data from all of the medical centers. Certain direct identifiers are removed, but dates of services are retained. The data set is distributed weekly to each UC Health site and will evolve over time as the UCHDW adds more patients and clinical depth.

UC CORDS, as well as the UCHDW, are using the OMOP Common Data Model version 5.3.1.

UC CORDS Data Description −



RE-COV-Ry

Real-world Evidence COVID-19 Registry

The Hackensack Meridian experience

Andrew Ip, MD, MSc

John Theurer Cancer Center
Hackensack Meridian Health
Division of Outcomes and Value Research

Hackensack Meridian *Health*

John Theurer Cancer Center

Consortium Member of
Georgetown | Lombardi
COMPREHENSIVE CANCER CENTER

NCI Comprehensive Cancer Center
A Cancer Center Designated by the National Cancer Institute

Developing RE-COV-Ry database

- Prospective, observational database of patients with COVID-19 at one of 13 Hackensack Meridian Health (HMH) hospitals spanning NJ

- **~5000** in current database with demographics, presenting features and labs on admission, labs on entry to ICU (if applicable), treatments (general) and survival outcomes

  - Use of **REDCap** (Research Electronic Data Capture) to capture, store, and export data

Developing RE-COV-Ry database

- Finding patients rapidly
  - EPIC EHR was used to run a report on all **POSITIVE** or **SUSPECTED** COVID19 patients (automatically flagged in patient's chart if test is positive or pending), and we abstracted based off of this report.
- No automated report tool available to abstract key outcomes, labs, treatments, demographics*
  - Data Managers (Research nurses) were used for **MANUAL abstraction** into REDCap, with supervision by investigators to ensure data quality
  - *demographics later were pulled automatically from EPIC electronic health records

Utilizing Real-World data to evaluate therapeutics

- Studied ~200 tocilizumab (IL-6 inhibitor) in an ICU cohort of ~700 patients hospitalized at HMH
- An adjusted cox proportional hazards regression model was used to estimate association of tocilizumab use and overall survival
- Data quality issues
  - **Missing** data (labs, treatments)
  - Time-dependent variables reviewed for accuracy (mechanical ventilation, time to treatment, admission or discharge dates)

Ip et al, under revision

Utilizing Real-World data to evaluate therapeutics

Image removed – data under revision

- An association of **improved OS** was seen in the tocilizumab group
- The ICU mortality rate in BOTH groups is quite high, reflecting the early epidemic in North Jersey
- Attempts to adjust for confounders , immortal time bias, indication bias, can be tedious

Ip et al, under revision
*data not to be shared without permission

Future Directions of RE-COV-Ry

- Ongoing analysis of other therapeutics
  - Outpatient Hydroxychloroquine (manuscript in preparation)
  - Empiric anticoagulation in ICU (analysis still pending)
  - Inpatient hydroxychloroquine (manuscript under review)
- Collaborations
  - w/ FDA Evidence Accelerator (COVID-19 Therapeutics) in partnership with COTA
  - w/ University of Miami to develop an in-patient risk score predicting intubation / mortality
  - w/ Center of Discovery and Innovation to connect viral genomic studies with database

RE-COV-Ry
Changes to Improve registry

- REDCap will integrate and automatically pull from our EHR demographic, clinical data including diagnoses, medications, laboratory (over 3,000 lab variables available)

    - ongoing testing to implement within HMH

- EPIC Business Intelligence to automate more data pulls on specific queries
    - Human resource fatigue / burnout from manual data pull

- Second REDCap administrator needed

Thank you

Acknowledgements:

Stuart Goldberg, MD (Director of Outcomes and Value Research)

Michael Marafelias (REDCap manager)

All nurses, physicians who abstracted data at John Theurer Cancer Center

Georgetown Statistical group (Jaeil Ahn, Shuqi Wang)

COTA for analytics support

Andrew.ip@hackensackmeridian.org

THE INNOVATION CENTER FOR BIOMEDICAL INFORMATICS

GEORGETOWN UNIVERSITY
Georgetown University Medical Center

COVID-19: Accelerating research, learning and improvement

Subha Madhavan, PhD, FACMI
Chief Data Scientist
Georgetown University Medical Center
@subhamadhavan
subha.madhavan@georgetown.edu

Council of Medical Specialty Societies
July 8, 2020
2 PM ET to 3.30 PM ET

What Academic Medical Centers are Facing

1. Data challenges

- Geographic and political divides
- Burden of emergency orders
- Socio-cultural, ethical, legal, trust issues
- Data collection, standardization, integration, reporting

2. Pathways forward

- Usecase-driven
- Reuse of existing data infrastructure to solve problems efficiently
- Training and education
- Convening and coordinating activities
- Research data networks
- Security, Privacy & Compliance

3. Example Projects

- First look at data from TERAVOLT registry
- AI approaches to organize massive research and scientific progress
- Immuno-genomic analysis of COVID-19 patients

Thoracic cancERs internAtional coVid 19 cOLlaboraTion TERAVOLT Registry

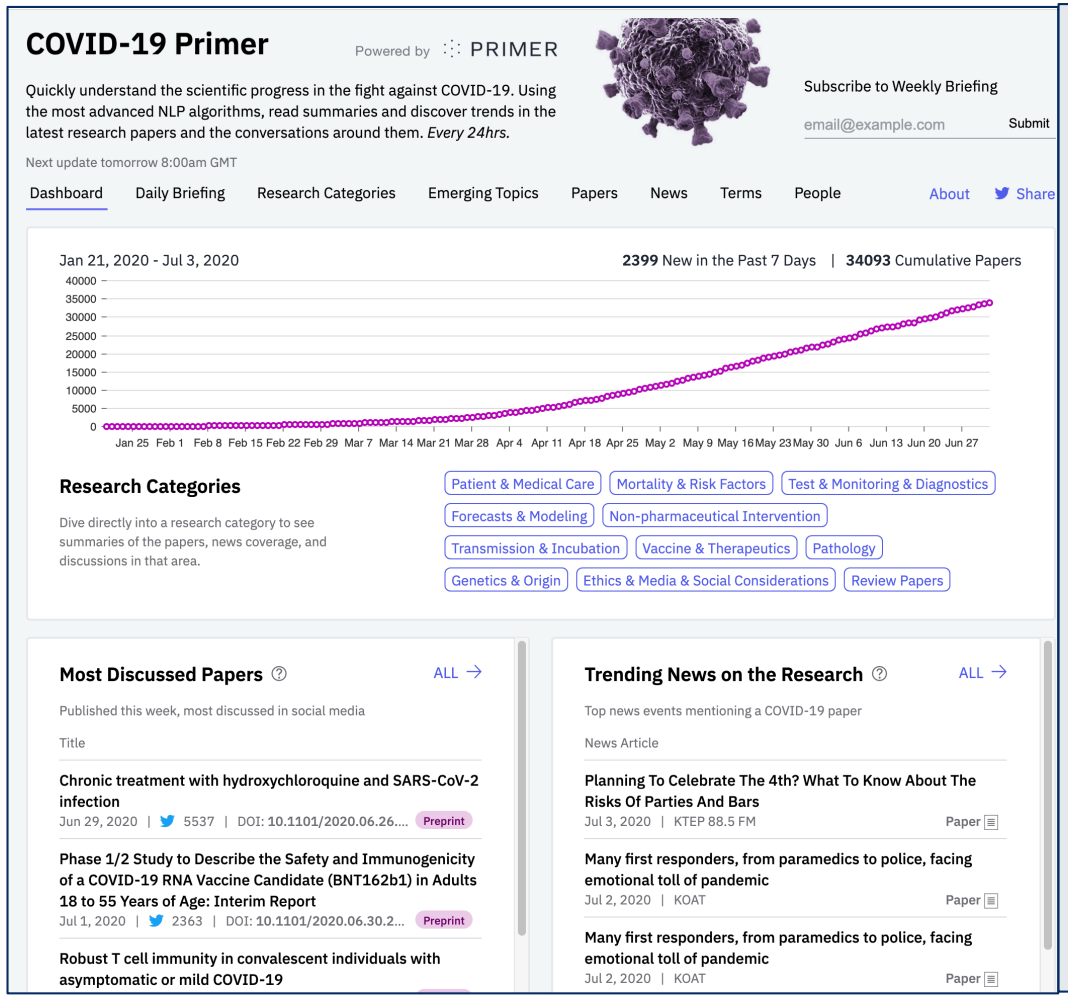
- Patients with COVID-19 and Thoracic Cancers
- 200 patients from 8 countries included
- **Time period:** March 26 – April 12
- **Data collected:** Demographics, Diagnostic test, Symptoms, Comorbidities, Concomitant medication, History of cancer, Complications, COVID-related treatments administered, Imaging modality, outcome (admission to ICU, death, reason for death, discharged, continue/delay of oncological treatment
- **Median age** was 68 years
- **ECOG** performance status: 0-1
- Stage: Majority (74%) had Stage 4 disease
- Majority were current or former smokers

Thoracic cancERs internAtional coVid 19 cOLlaboraTion TERAVOLT Registry

- 76% were hospitalized
- 33% died
  - 79% of deaths were due to COVID-related complications
- Only 10% met criteria and were admitted to the ICU
- Following factors were associated with higher risk of death
  - Being older than 65 years
  - Being a current/former smoker
  - Currently receiving chemotherapy treatment
  - Presence of comorbidities
- Whether mortality could be reduced with treatment in intensive care remains to be determined

COVID-19 Flattening the Curve Data Visualization Challenge

COVID-19 Flattening the Curve Data Visualization Challenge - 1st Prize



- Aggregation and synthesis of papers, news articles and social media posts
- Use of advanced NLP algorithms for topic modeling, relevance and current trends
- Updated every 24 hours
- Weekly briefings

New One-year Masters Program in Health Informatics & Data Science

GEORGETOWN UNIVERSITY
Biomedical Graduate Education

icbi
THE INNOVATION CENTER FOR BIOMEDICAL INFORMATICS
OFFICE OF HEALTH INFORMATICS AND DATA SCIENCE

Master's Degree
Health Informatics & Data Science

The Master's in Health Informatics & Data Science (HIDS) at Georgetown University is an industry-driven, career-ready program focused on current and emerging technologies that will inform healthcare. The program offers students access to Georgetown's distinguished faculty and unique opportunities in the nation's capital. Students will gain core competency in data science, big data analytics, digital ethics, artificial intelligence and machine learning applications to inform health outcomes and reduce costs. Our goal is to develop responsible leaders in health data science and medical informatics with rewarding careers in industry, government or academia.

For more information, visit
https://healthinformatics.georgetown.edu/

PROGRAM FEATURES

12 month
accelerated program

32 credits
Full-time or part-time

Final capstone project with
industry partners or government agencies

PROGRAM LEADERSHIP

Subha Madhavan, Ph.D., FACMI
Director, M.S HIDS

Adil Alaoui, M.S, MBA.
Co-Director, MS HIDS
Director, Industry Capstone Program

Yuriy Gusev, Ph.D.
Curriculum Director

Shruti Rao, M.S, MBA
Program Coordinator

WE ARE NOW ACCEPTING APPLICATIONS FOR FALL 2019 ADMISSION

Application Deadline: July 1st, 2019

For further information visit:

Website: https://healthinformatics.georgetown.edu/
Email: hids@georgetown.edu
Phone: +1 (202) 687 1093

GEORGETOWN UNIVERSITY
Biomedical Graduate Education

icbi
THE INNOVATION CENTER FOR BIOMEDICAL INFORMATICS
OFFICE OF HEALTH INFORMATICS AND DATA SCIENCE

Master's Degree
Health Informatics & Data Science

https://healthinformatics.georgetown.edu/

The program includes comprehensive coursework, hands-on projects and a mandatory **Capstone Project** that is the culmination of the student's coursework and experience at Georgetown University. Students will work with a host organization and an academic mentor to critically analyze a real world problem, design, conceptualize and implement a solution that applies the knowledge and skills acquired during the Masters program.

FALL

- Introduction to Health Data Science & Analytics (HIDS-501)
- Utilizing Data in Electronic Medical Records (HIDS-502)
- Precision Health Informatics (HIDS-503)
- Evidence based data analysis in population health (HIDS-504)

SPRING

- Massive Data Fundamentals (ANLY-502)
- Advanced topics in Health Informatics (HIDS-505)
- AI for health applications (HIDS-506)
- Digital Health Applications (HIDS-507)

SUMMER

- Human Factor Engineering, Usability and Safety (HIDS-508)
- Imaging Informatics (HIDS-509)
- Mandatory Capstone (HIDS-510)

INDUSTRY & GOVERNMENT ADVISORY PANEL

- Sage Bionetworks
- ESAC
- FDA
- Deloitte
- Flatiron Health
- COTA
- Syapse
- TrinetX
- INOVA
- MedStar Institute for Innovation (MI2)
- Cerner Corporation
- Indivumed Group
- Argentys
- CVS Health
- PwC
- Open Source Electronic Health Record Alliance (OSEHRA)
- RadAmerica
- LexisNexis
- Assymetrik

APPLICATION CHECKLIST

- 4-year Bachelor's degree
- Grade point average of 3.0 or higher
- Official transcripts
- 2 letters of recommendation
- Academic Statement of Purpose
- Interview with faculty
- TOEFL/IELTS Score (if applicable)
- No GRE

Thank you

- http://icbi.georgetown.edu/
- subha.madhavan@georgetown.edu
- @subhamadhavan





Automating statewide collection of medical surge capacity during a pandemic

July 8, 2020

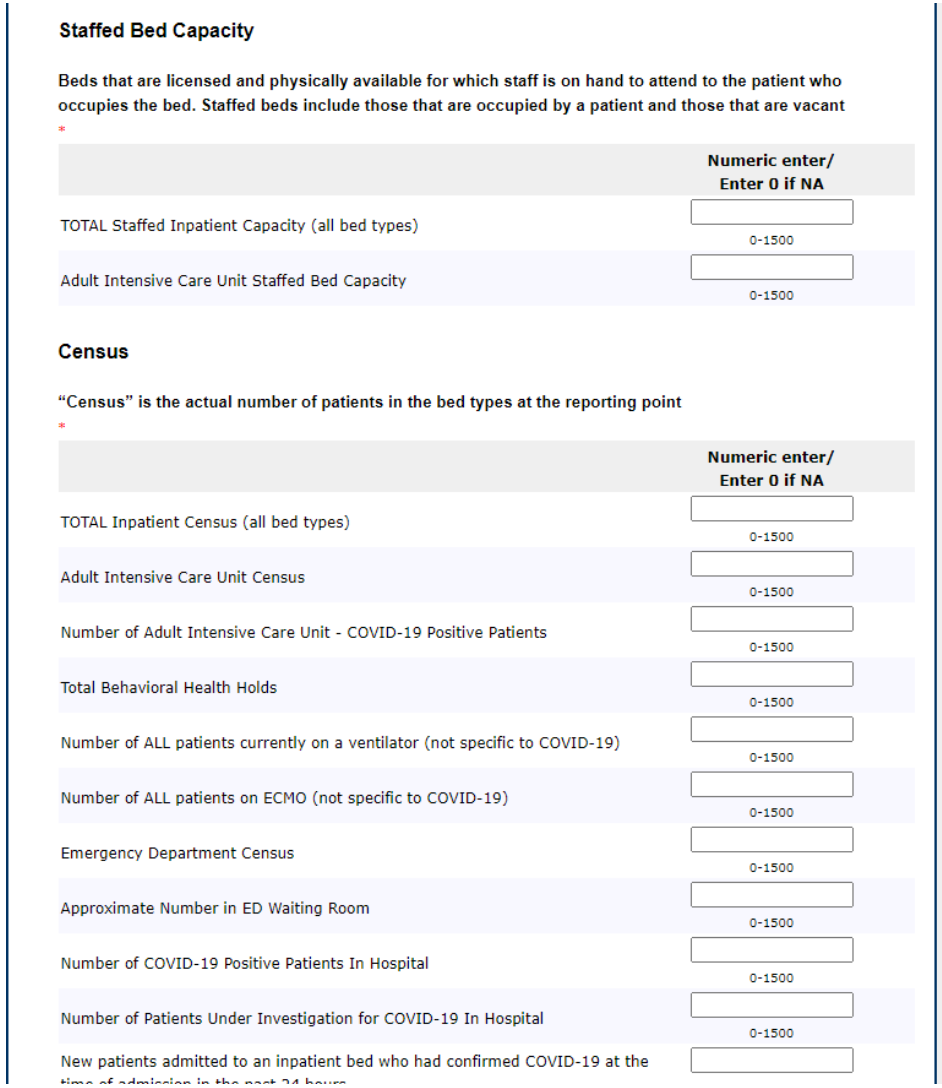
Jessie Tenenbaum, PhD, FACMI
Chief Data Officer, NC DHHS
Duke University School of Medicine
@jessiet1023

Medical Surge Data: “Stuff, Staff, Space”

- How many beds are full, how many available?
- ICU beds?
- Ventilators?
- How many COVID+ in the hospital? In ICU? Admitted in past 24 hours?
- Doctors, nurses?
- Other questions- ED waiting room, ECMO, labs, morgue, etc.

MedSurge Survey

- Daily email to ~120 hospitals statewide
- Person filling out survey rotates
- Data validation is difficult
- Results highly visible
- Missing hospitals
- 1+ hour nightly data cleaning



Staffed Bed Capacity

Beds that are licensed and physically available for which staff is on hand to attend to the patient who occupies the bed. Staffed beds include those that are occupied by a patient and those that are vacant

*

	Numeric enter/ Enter 0 if NA
TOTAL Staffed Inpatient Capacity (all bed types)	0-1500
Adult Intensive Care Unit Staffed Bed Capacity	0-1500

Census

"Census" is the actual number of patients in the bed types at the reporting point

*

	Numeric enter/ Enter 0 if NA
TOTAL Inpatient Census (all bed types)	0-1500
Adult Intensive Care Unit Census	0-1500
Number of Adult Intensive Care Unit - COVID-19 Positive Patients	0-1500
Total Behavioral Health Holds	0-1500
Number of ALL patients currently on a ventilator (not specific to COVID-19)	0-1500
Number of ALL patients on ECMO (not specific to COVID-19)	0-1500
Emergency Department Census	0-1500
Approximate Number in ED Waiting Room	0-1500
Number of COVID-19 Positive Patients In Hospital	0-1500
Number of Patients Under Investigation for COVID-19 In Hospital	0-1500
New patients admitted to an inpatient bed who had confirmed COVID-19 at the	

National Healthcare Safety Network (NHSN) Hospital capacity module data collection instructions

Instructions for Completion of the COVID-19 Patient Impact and Hospital Capacity Module Form (CDC 57.130)

Data Field	Instruction for Data Collection
Facility ID #	The NHSN-assigned facility ID will be auto-entered by the computer.
Summary Census ID #	Auto-generated by the computer.
Date for which patient impact and hospital capacity counts are reported	*Required*. Select the date for which the recorded data was collected for the following questions.

Section-1: Patient Impact

Data Field	Instruction for Data Collection
HOSPITALIZED: Patients currently hospitalized in an inpatient bed who have suspected or confirmed COVID-19	Enter the number of patients hospitalized in an inpatient bed at the time the data is collected who have suspected or confirmed COVID-19. This includes the patients with laboratory-confirmed or clinically diagnosed COVID-19. **Confirmed:** A patient with a laboratory-confirmed COVID-19 diagnosis **Suspected:** A patient without a laboratory confirmed COVID-19 diagnosis who, in accordance with CDC's Interim Public Health Guidance for Evaluating Persons Under Investigation (PUIs), has signs and symptoms compatible with COVID-19 (most patients with confirmed COVID-19 have developed fever and/or symptoms of acute respiratory illness, such as cough, shortness of breath or myalgia/fatigue).
HOSPITALIZED and VENTILATED: Patients currently hospitalized in an inpatient bed who have suspected or confirmed COVID-19 and are on a mechanical ventilator	Enter the number of patients hospitalized in an inpatient bed who have suspected or confirmed COVID-19 and are currently on a mechanical ventilator* at the time the data is collected . This includes the patients with laboratory-confirmed or clinically diagnosed COVID-19.
	*Ventilator: Any device used to support, assist or control respiration (inclusive of the weaning period) through the application of positive pressure to the airway when delivered via an artificial airway, specifically an oral/nasal endotracheal or tracheostomy tube. Note: Ventilation and lung expansion devices that deliver positive pressure to the airway (for example: CPAP, BiPAP, bi-level, IPPB and PEEP) via non-invasive means (for example: nasal prongs, nasal mask, full face mask, total mask, etc.) are not considered ventilators unless positive pressure is delivered via an artificial airway (oral/nasal endotracheal or tracheostomy tube).
HOSPITAL ONSET: Patients currently hospitalized in an inpatient bed with onset of suspected or confirmed COVID-19 fourteen or more days after hospital admission due to a condition other than COVID-19	Enter the number of patients hospitalized in an inpatient bed at the time the data is collected with onset of suspected or confirmed COVID-19 fourteen or more days after hospitalization (admission date = hospital day 1). This includes laboratory-confirmed or clinically diagnosed COVID-19 cases.
ED/OVERFLOW: Patients with suspected or confirmed COVID-19 who are currently in the Emergency Department (ED) or any overflow location awaiting an inpatient bed	Enter the number of patients with suspected or confirmed COVID-19 who are in the Emergency Department(ED) or any overflow/expansion location awaiting placement in an inpatient bed at the time the data is collected . This includes patients with laboratory-confirmed or clinically diagnosed COVID-19. Overflow locations include any physical locations created to accommodate patients include but not limited to 24-hour observation units, hallways, parking lots, or tents.
ED/OVERFLOW and VENTILATED: Patients with suspected or confirmed COVID-19 who currently are in the ED or any overflow location awaiting an inpatient bed and on a mechanical ventilator	Enter the number of patients with suspected or confirmed COVID-19 who are in the ED or any overflow/expansion location on a mechanical ventilator* at the time the data is collected . This includes patients with laboratory-confirmed or clinically diagnosed COVID-19.

Enter Appriss/Open Beds

- Controlled substance use disorder platform- used in NC
- Purchased Open Beds

- Cloud-based platform that tracks behavioral health resources

Repurpose for Covid resources!

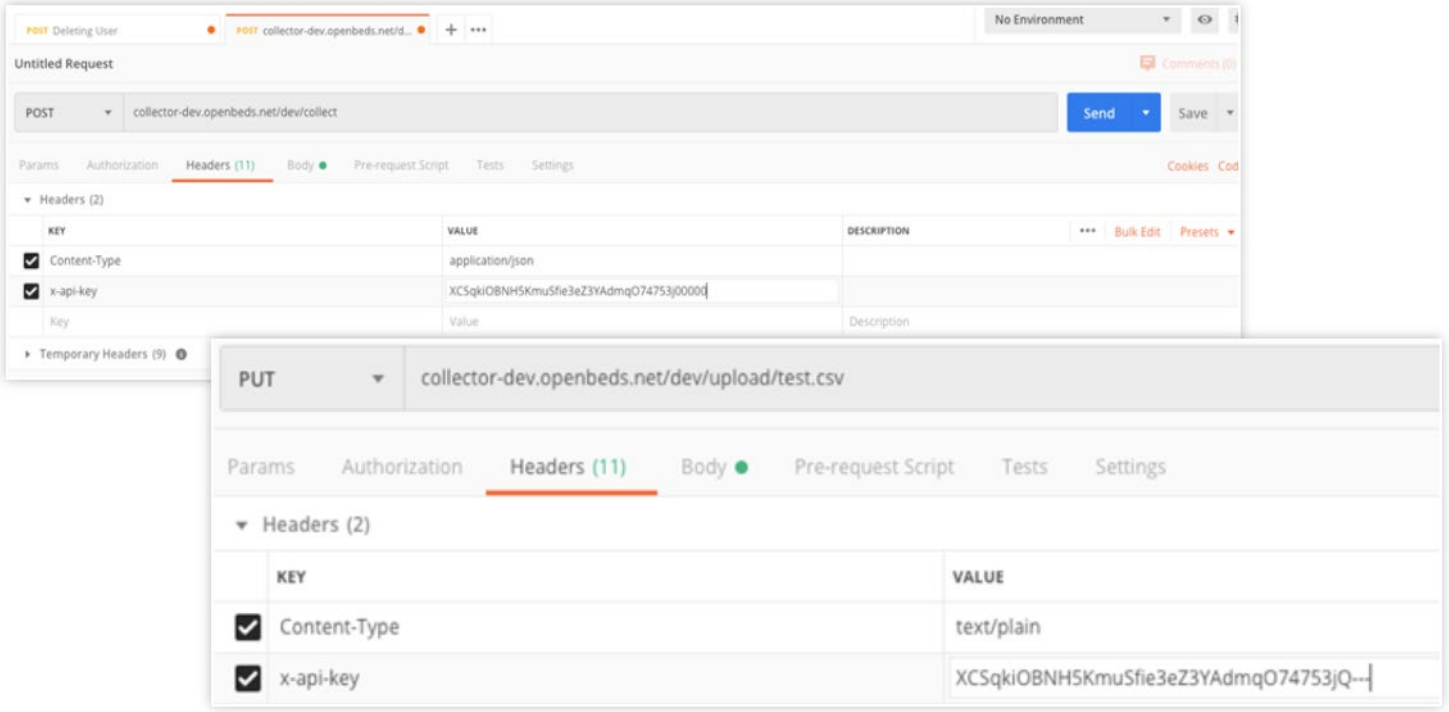
MedSurge Data Automation

- Appriss developed API to ingest resource use from HL7 feeds
- Hospital systems onboarded to extract required data elements, feed into Appriss system
- Appriss pushes data hourly to NCDHHS
- As data feeds go live, hospitals can stop filling out manual survey- still fulfill state and federal reporting requirements

Dashboard

Resource Availability – Submitting Data

1) Automated API



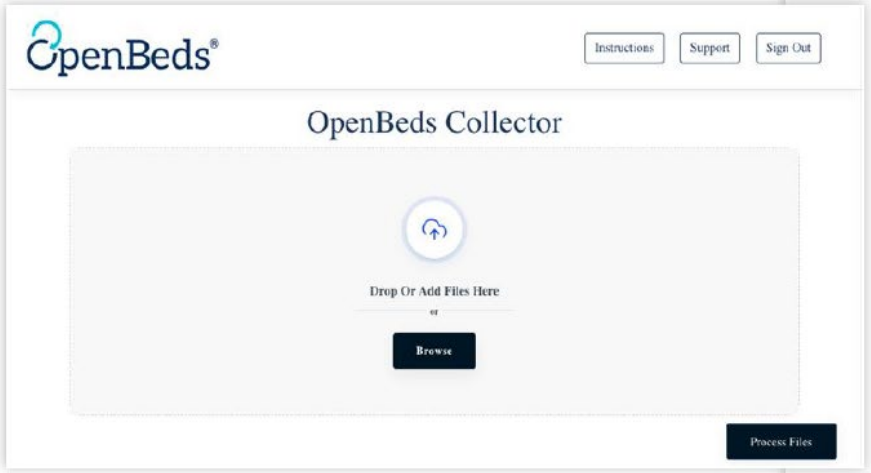
2) Automated CSV

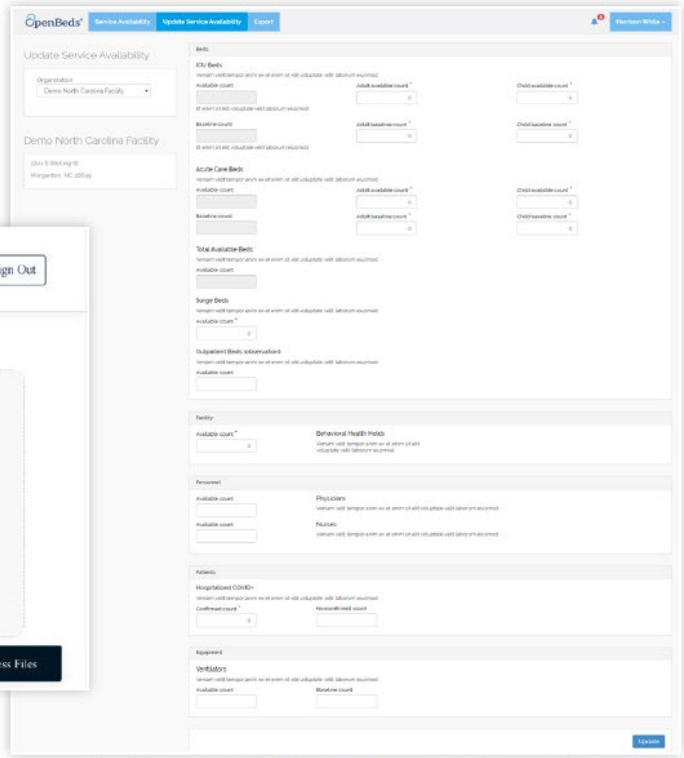
> Hospitals have the flexibility to submit data via four mechanisms

Dashboard

Resource Availability – Submitting Data

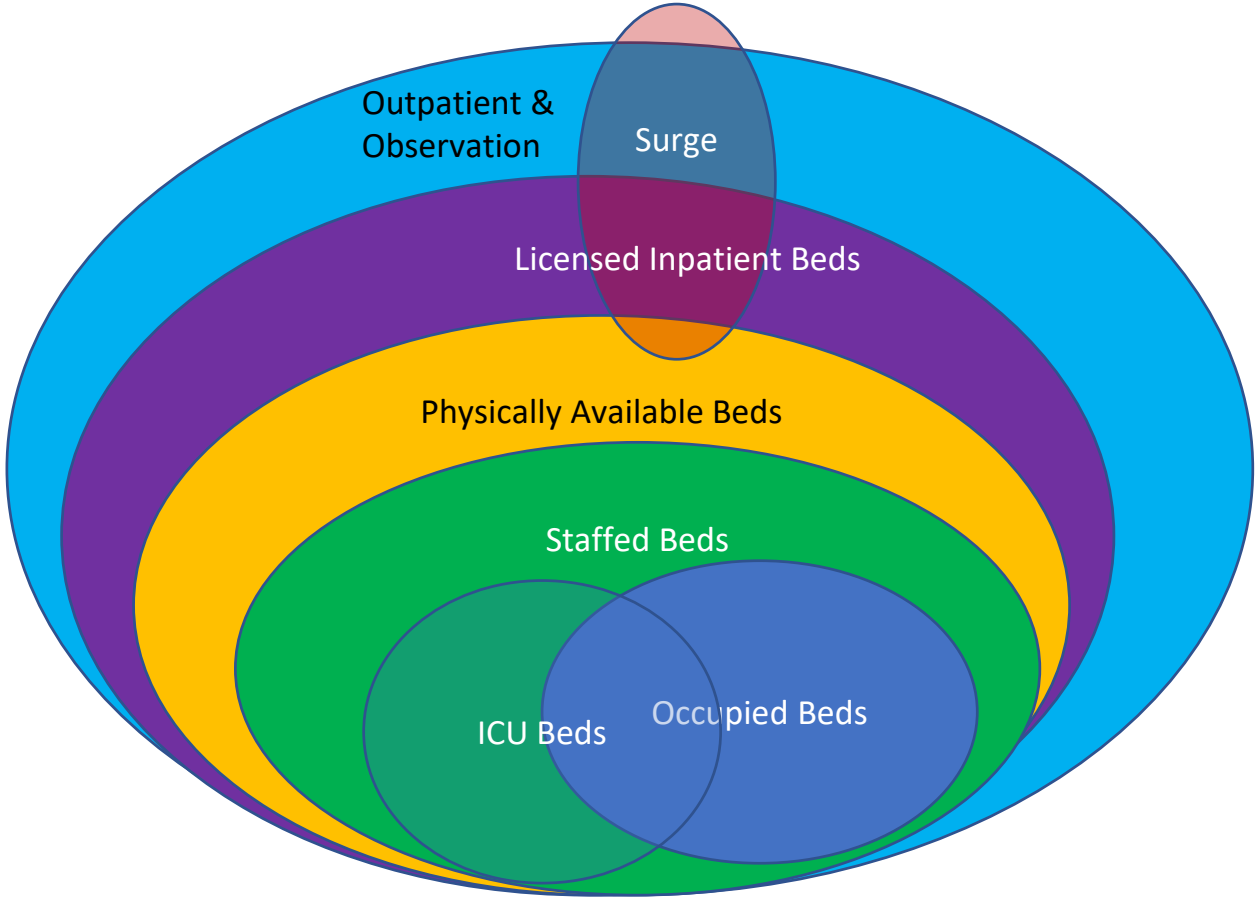
3) Manual CSV



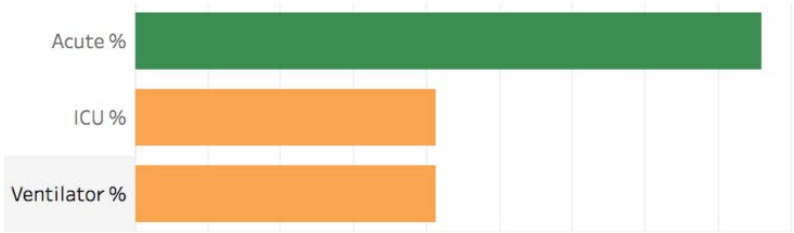


4) Manual direct entry into form at Tracker

A bed by every other name…

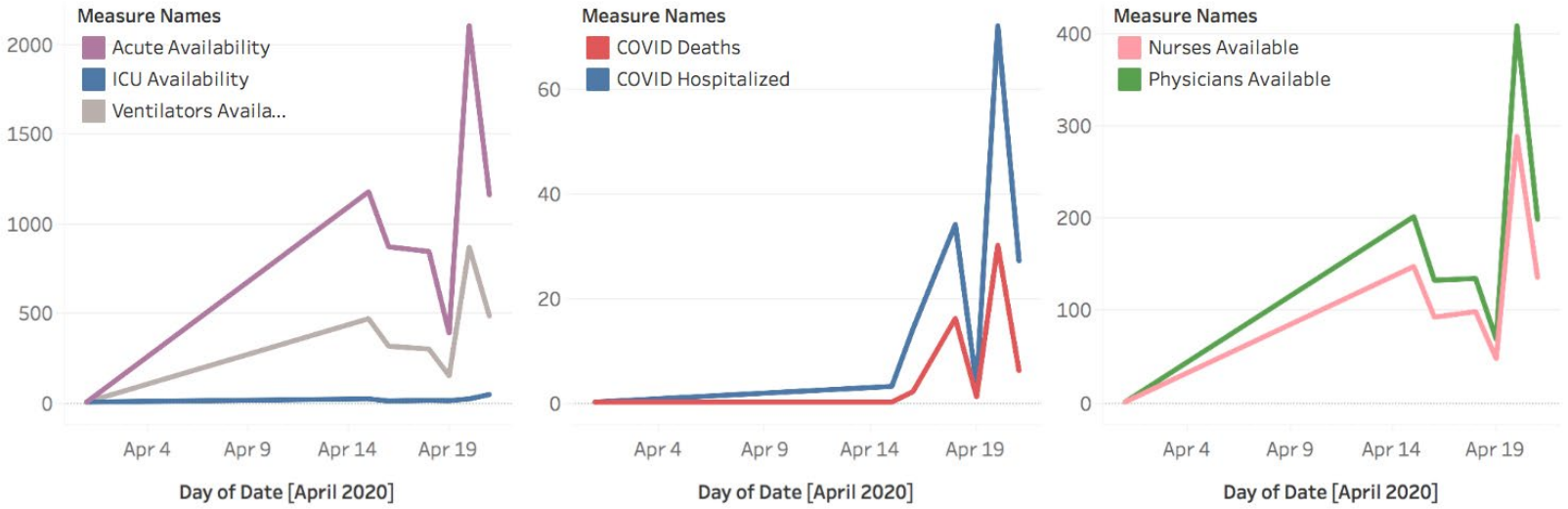


Summary and Metrics for
Mission Hospital McDowell (FKA McDowell Hospital)



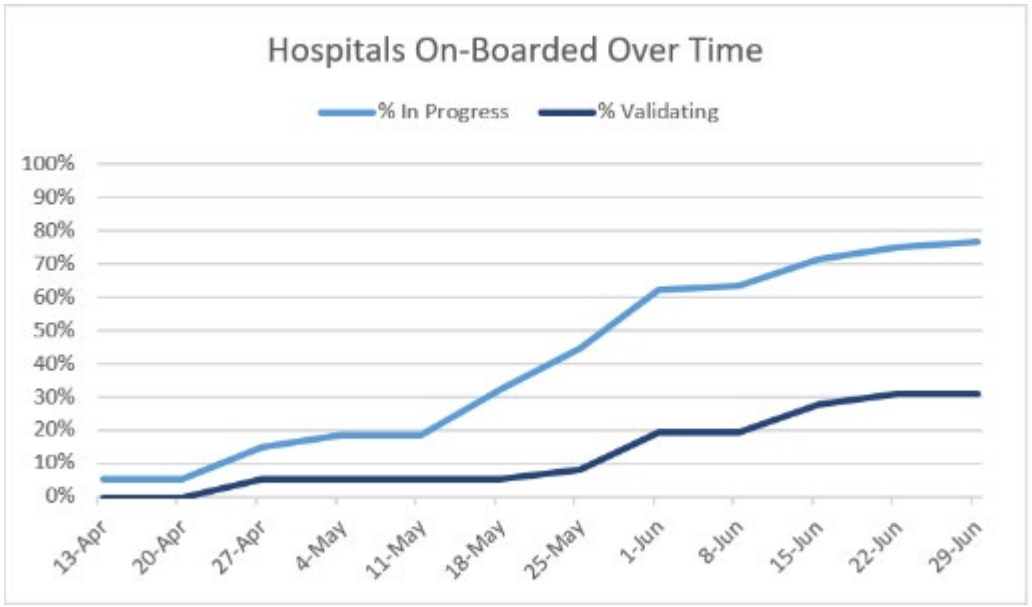
Hospital Metrics

organization	Acute Availability	ICU Availability	Total Availability	Ventilators Available	COVID Deaths	Confirmed COVID	Nurses Available	Physicians Available	Bed Surge
Mission Hospital McDowell (..	6,519	102	6,621	2,564	55	98	807	1,141	189



Hospital Onboarding

Onboarding Status of Hospitals		
Total Done	35	31%
Currently In Progress	53	47%
Currently Not Started	24	21%



Challenges

- Surprisingly hard to settle on data element definitions, even when you think you understand them
- Tension between complying with what's written and what we were told is sufficient (and not overburdening hospital staff)
- Onboarding systems over time- merging data from 2 different sources
- Some data elements were dropped- not worth the squeeze (and priorities change)

Summary

- Successful public/private partnership between NC DHHS and Appriss Inc.
- NC hospital personnel can spend their time on patient care, not data entry
- State gets accurate, up-to-date snapshot of available resources
- Saves resources on state side as well- eliminate need for manual data cleaning

Acknowledgements

NC DHHS

- **Charles Carter**
- **Danielle Brady**
- Michelle Kish
- Kimberly Clement
- Jean Chiang

Mission Health

- John Brown
- Cameron Hurst

Appriss

- **Nishi Rawat**
- **Lauren Whitsell**
- Bruce Bridges
- Harrison White

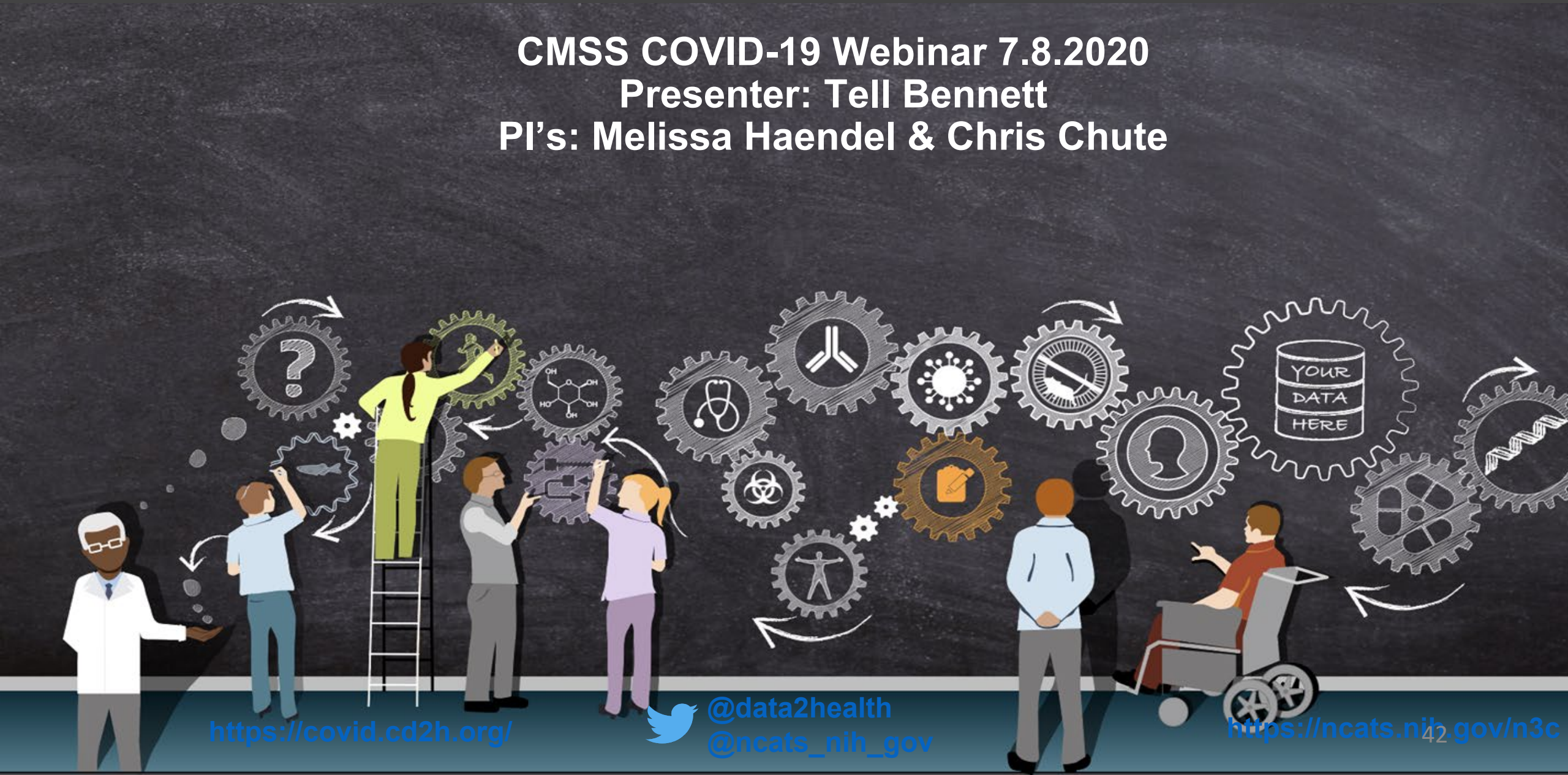
UNC

- Mike Plesh

And many, many others

National COVID Cohort Collaborative (N3C)

CMSS COVID-19 Webinar 7.8.2020
Presenter: Tell Bennett
PI's: Melissa Haendel & Chris Chute



https://covid.cd2h.org/

@data2health
@ncats_nih_gov

https://ncats.nih.gov/n3c

National COVID Cohort Collaborative

This pandemic highlights urgent needs

- ML algorithms (diagnosis, triage, predictive, treatment pathways, etc.)
- Best practices for resource allocation
- Drug discovery
- Coordinate our efforts to maximize efficiency

All these things require the creation of a comprehensive clinical data set

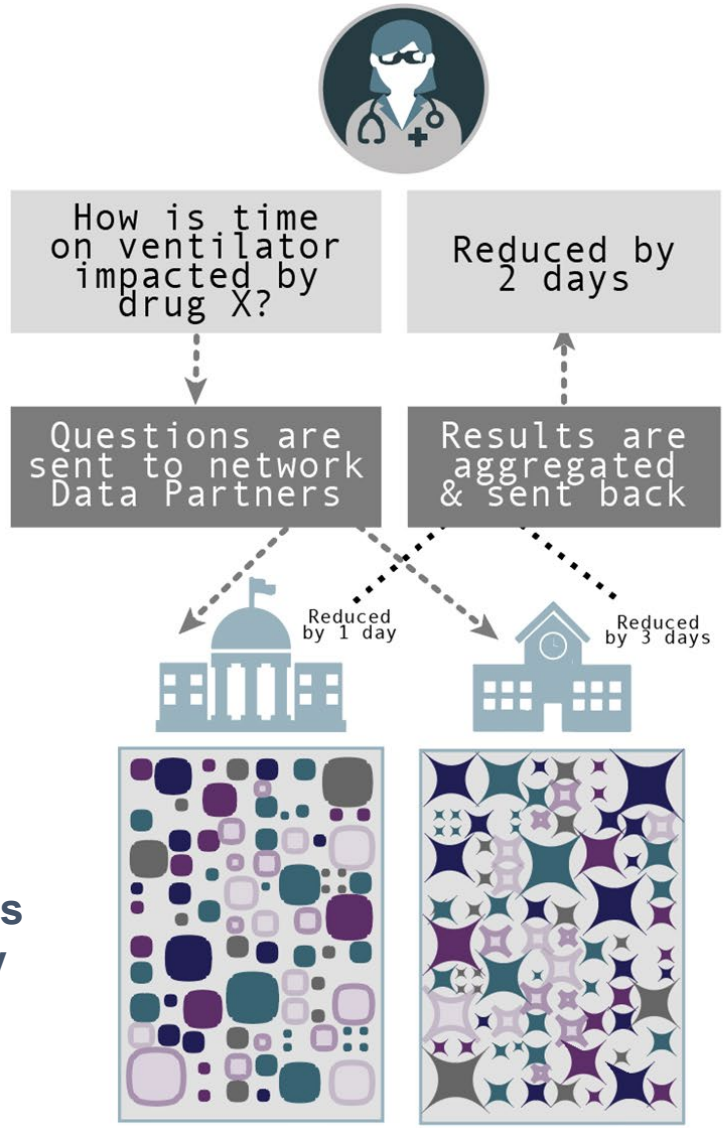
National COVID Cohort Collaborative

But, am I not already sending data?

N3C is synergistic with distributed data networks!

Centralizing patient-level data makes it possible to ask qualitatively different and more powerful questions, but is only possible due to each institution having their data in a common data model.

Federated querying



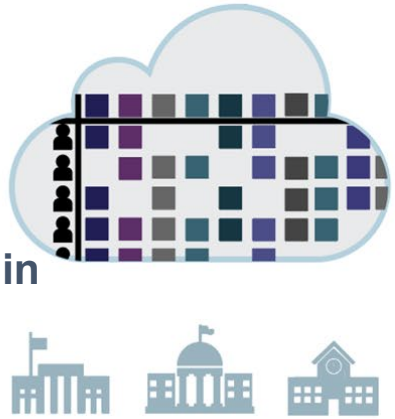
Data resides locally

Centralized analytics

In patients under age 60, which factors are most predictive of severe outcomes?

Collaboratively build, test, and refine algorithmic classifiers

Identify novel associations

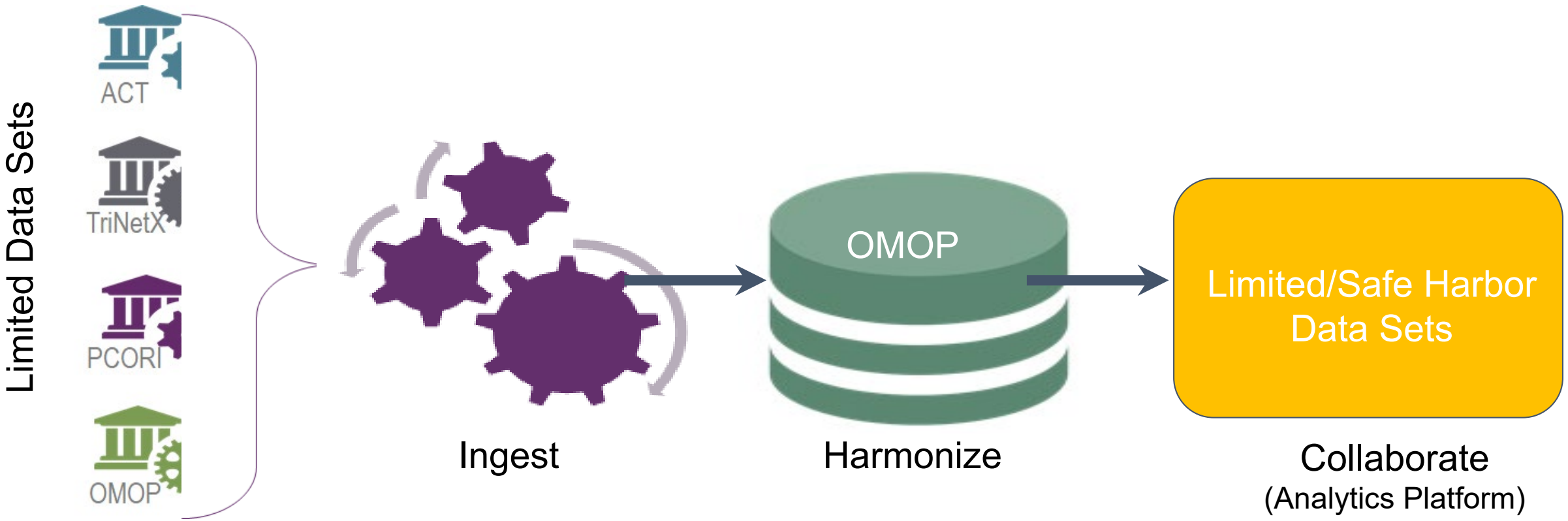


Data resides centrally in a secure enclave

N3C cloud

National COVID Cohort Collaborative

N3C Overview



1 Data partnership & governance

2 Phenotype & Data acquisition

2 Data ingest & harmonization

3 Collaborative analytics & FAIR Sharing/Credit

National COVID Cohort Collaborative

Timeline

AMIA Kickoff
APRIL 13

Workstreams launched
JHU IRB established

APRIL 20

APRIL 27

DTA finalized by NIH
Phenotype v1.0 published

N3C platform provisioned in Cloud
1st DTA signed
Platform training initiated

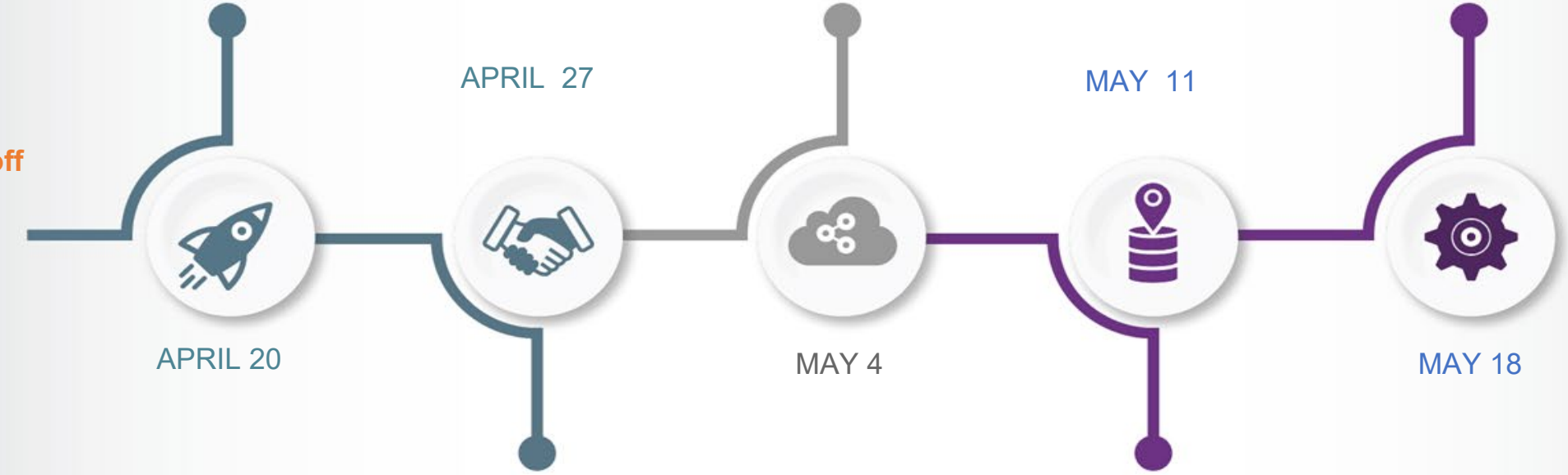
MAY 4

MAY 11

NCATS C.Austin community mtg
Phenotype & extract scripts published
Data Harmonization maps created

500+ members
71 requested DTAs; **37** signed as of 6/10
5 sites submitted data
49 people trained
1st Test ML models: intubation & AKI
Data Harmonization pipeline built

MAY 18



7/8/2020
49 DTAs executed
27 IRB protocols approved (23 reliance, 4 local)
24 Regulatory complete (both DTA and IRB)
36 Met with Data Acquisition Group
.....9 Deposited data:
.....4 - PCORI
.....3 - OMOP
.....1 - TriNetX
.....1 - ACT

National COVID Cohort Collaborative

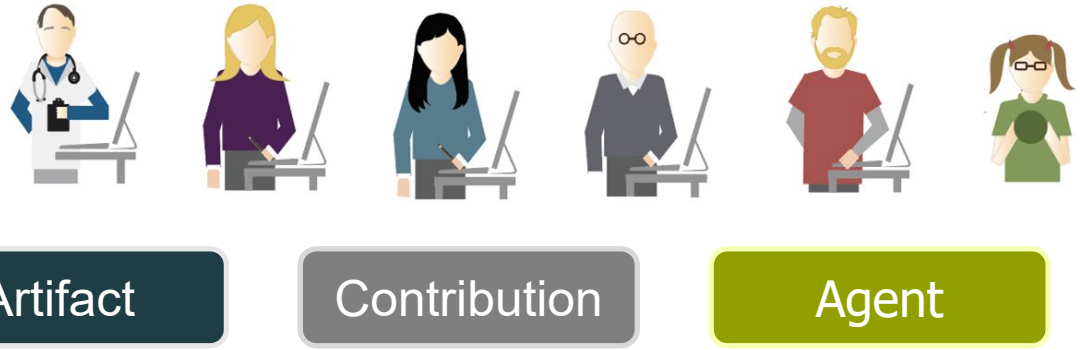
Data Partnership and Governance

Goal of the Data Use Agreement is broad access:

- COVID-Related research only
- **Open platform to all Credentialed researchers**
- Security: Activities in the N3C Enclave are recorded and can be audited
- Disclosure of research results to the N3C Enclave for the public good
- Analytics provenance
- Contributor Attribution tracking
- No download of data

Data Access Committee: [in formation]

Central IRB option through SMARTIRB



cd2h.org/attribution

National COVID Cohort Collaborative

Phenotype & Acquisition

Dual-purpose workstream:

1. Work with the community to write and maintain a computable phenotype for COVID-19.
2. Write and maintain a series of scripts to execute the computable phenotype in each of four common data models (CDMs): OMOP, i2b2/ACT, PCORnet, and TriNetX.

What does it look like to run our process locally?

1. **Run our phenotype code to define your COVID-19 cohort.**
   - 2x per week if possible
   - Code available for all data models, multiple database systems

2. **Run our lightweight local data quality checks.**
   - Checks only for "showstopper" issues to prevent back-and-forth after submission.

3. **Run our extract code, which will dump out data for that cohort to a series of flat files.**
   - Export code available as a Python script or plain SQL files.

4. **Zip up the flat files and transmit to N3C.**
   - Transmit via SFTP
   - Data will be picked up by the Data Harmonization team for integration into repository.

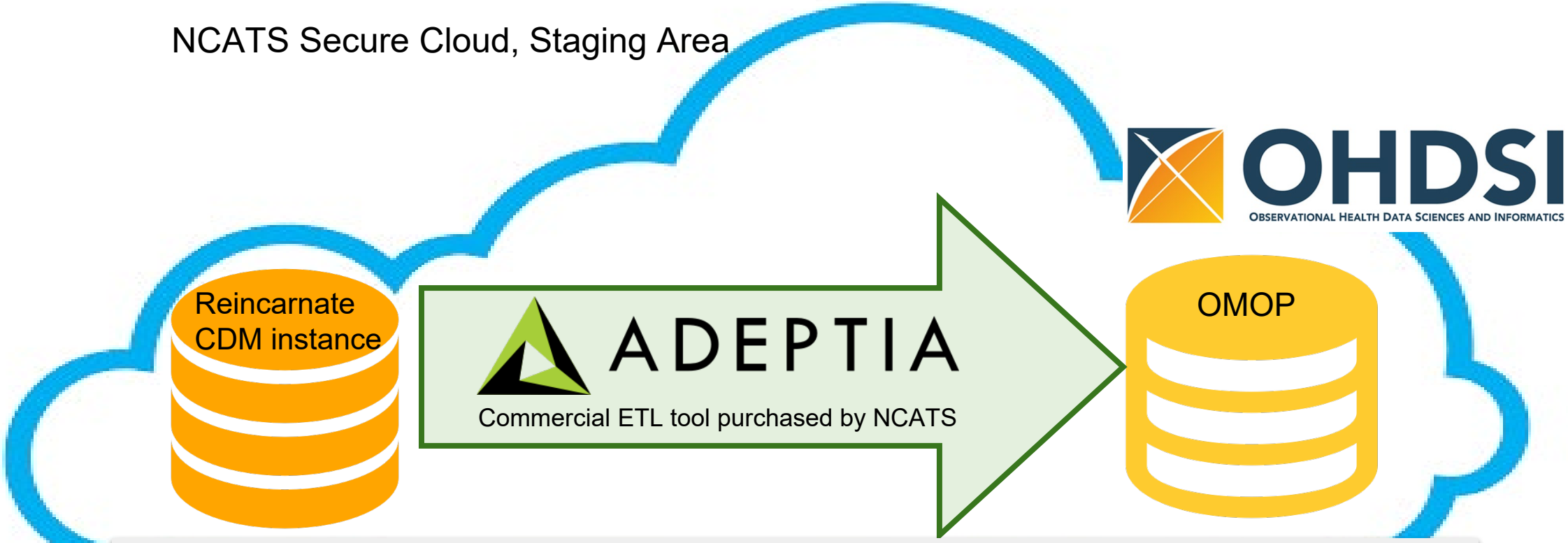
Support is available for all parts of this process!

Latest phenotype: covid.cd2h.org/phenotype

Documentation: covid.cd2h.org/phenotype-wiki

All specifications and software shared on GitHub

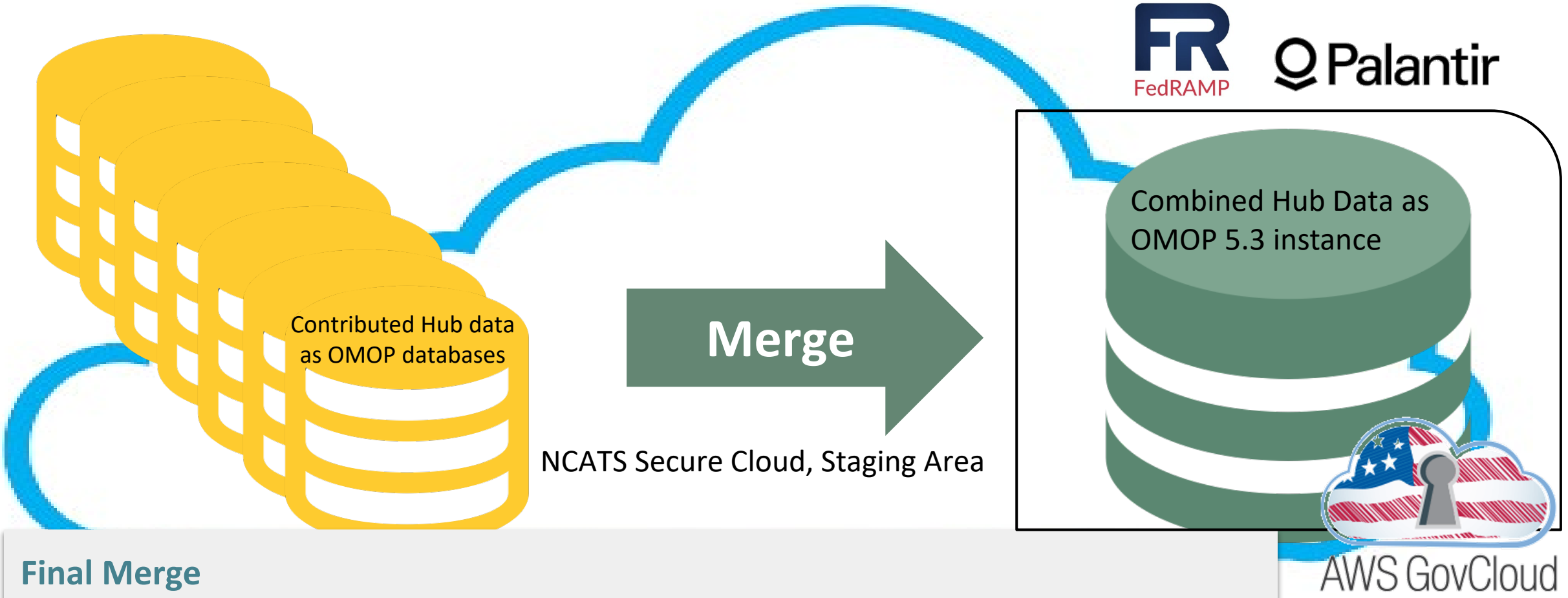
Data Harmonization: Transformation



Second Stage Ingestion

- Repair or encode aberrant data (COVID LOINC codes) ✔
- Transform source CDM into OMOP 5.3 ✔
- Leverage library of validated CDM to OMOP maps

Data Harmonization: Secure Integration

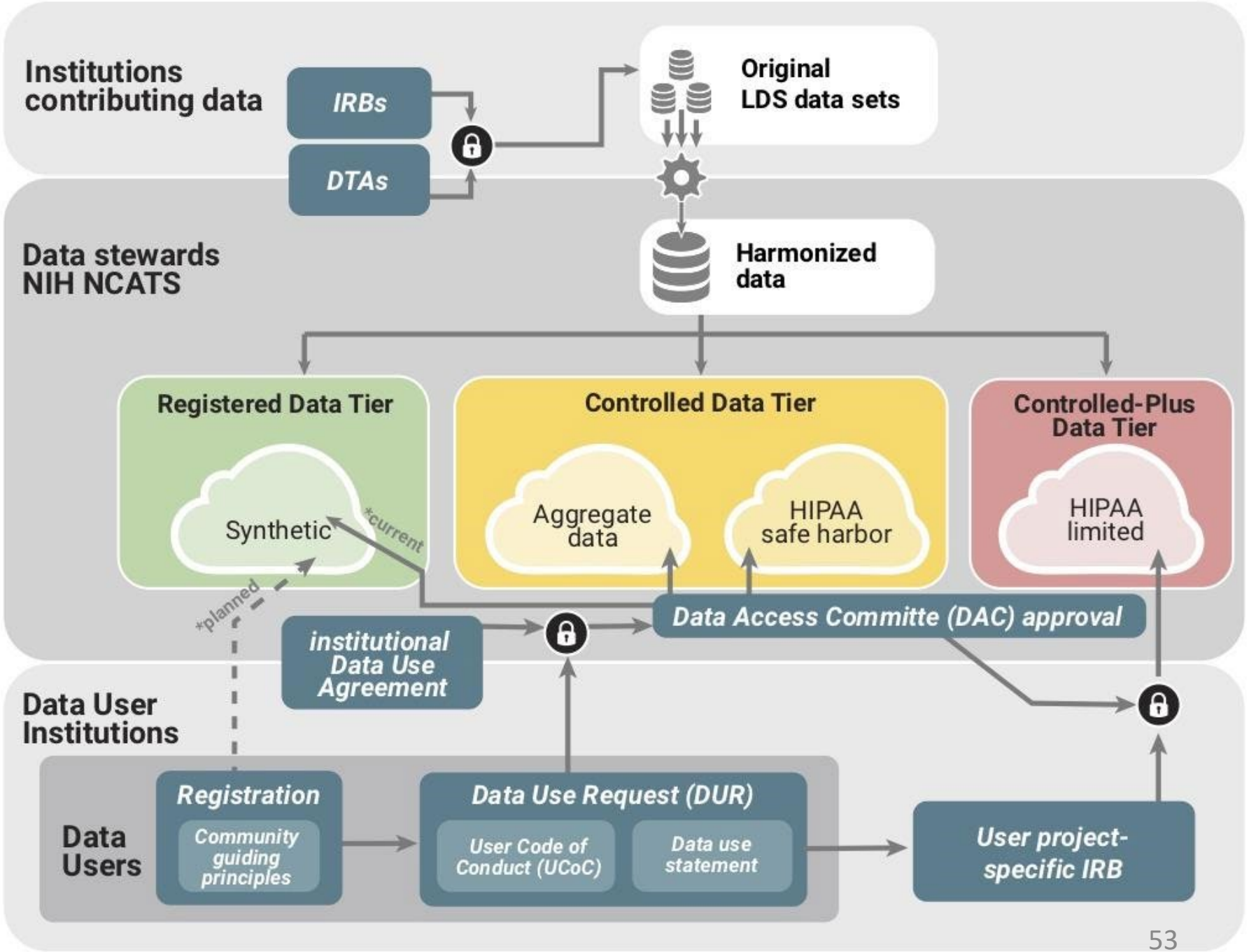


Final Merge

- OMOP versioned data from all sources combined into analytic database
- Analytic database will transfer to Palantir Analytic Platform

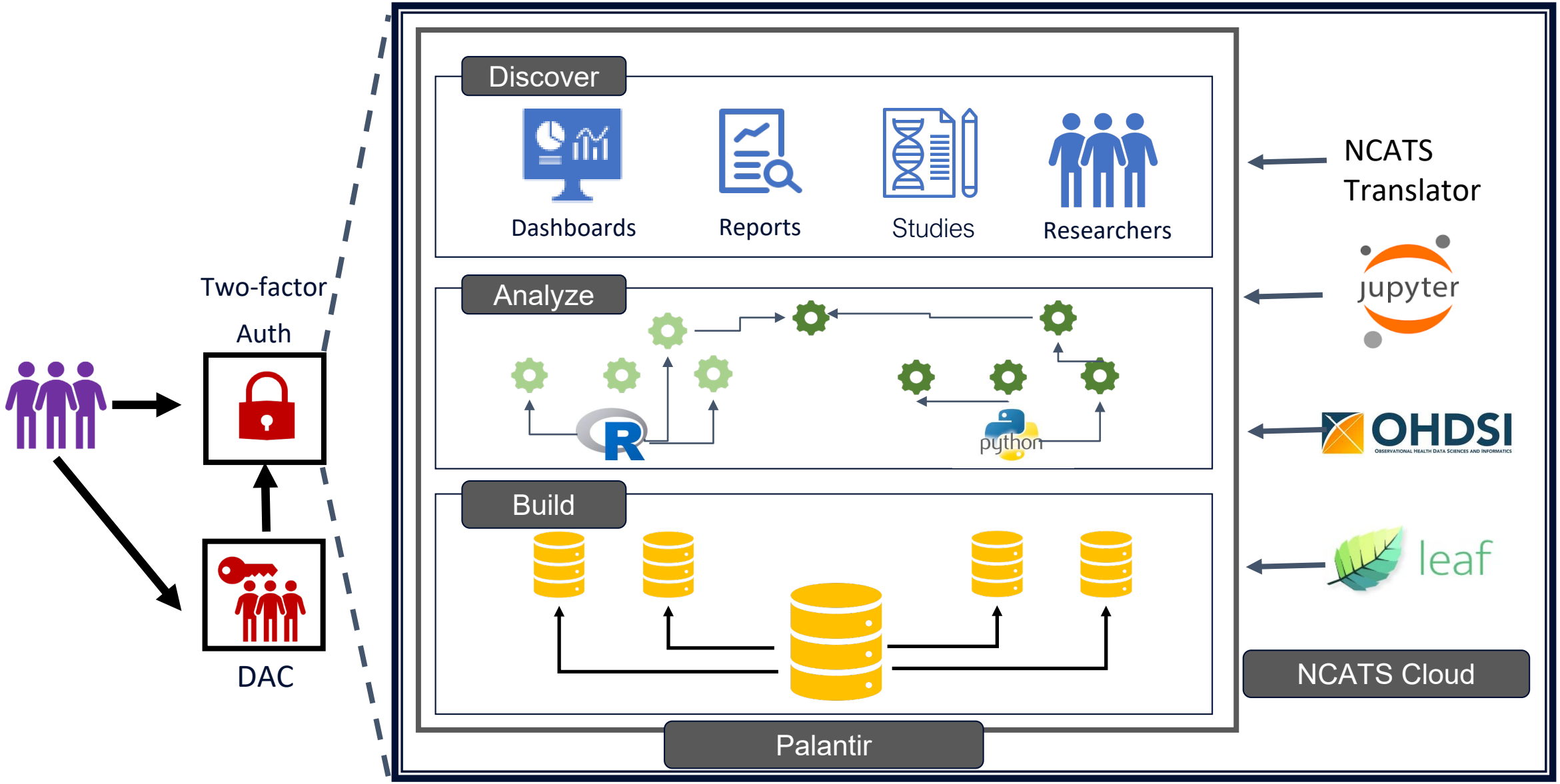
National COVID Cohort Collaborative

Data Access



National COVID Cohort Collaborative

N3C Secure Data Enclave



National COVID Cohort Collaborative

N3C Community Workstreams

NCATS N3C website: ncats.nih.gov/n3c

CD2H N3C website: covid.cd2h.org

Hub Partnership packet: https://covid.cd2h.org/partnership_welcome_packet

Onboarding to N3C: bit.ly/cd2h-onboarding-form

NATIONAL CENTER FOR DATA TO HEALTH

NIH National Center for Advancing Translational Sciences

National COVID Cohort Collaborative

National COVID Cohort Collaborative

Partners, Teams, Collaborators

NCATS
Chris Austin
Joni Rutter
Mike Kurilla
Clare Schmitt
Ken Gersing
Xinzhi Zhang
Erica Rosemond
Sam Bozzette
Lili Portilla
Chris Dillon
Penny Burgoon
Emily Marti
Meredith Temple-O'Connor
Sam Jonson
Christine Cutillo
Nicole Garbarini

NIH & HHS Partners
NCI
Janelle Cortner
Stephen Hewitt
Denise Warzel

FDA
Mitra Rocca
Scott Gideon
Wei Chen

NIDDK
Robert Star

NIGMS
Ming Lee

NCATS ITRB
Sam Michael
Mariam Deacy
Gary Berkson
Josephine Kennedy
Usman Sheikh
Mark Backus
Nam Ngo
Amit Virakatmath
Keats Kirsch
Sulochana Nunna
Rafael Fuentes
Reid Simon
Biju Mathew
Tim Mierzwa
Ke Wang
Kalle Virtaneva

CD2H
OHSU/OSU
Melissa Haendel
Anita Walden
Julie McMurry
Moni Munoz-Torres
Andrea Volz
Connor Cook
Racquel Dietz
Andrew Neumann
Rich Lorimor

Sage Bionetworks
Justin Guinney
James Eddy

U of Iowa:
Dave Eichmann
Alexis Graves

Northwestern:
Kristi Holmes
Justin Starren
Lisa O'Keefe

Washington U.
Philip Payne
Albert Lai
Tom Dillon

CD2H
U. Of Washington
Adam Wilcox
Liz Zampino

Johns Hopkins U
Chris Chute
Tricia Francis

Jax Labs
Peter Robinson

Scripps
Chunlei Wu

Teams
Governance
Sage Bionetworks
John Wilbanks
Christine Suver

Data Harmonization
JHU
Davera Gabriel
Stephanie Hong
Harold Lehmann
Tanner Zhang
Richard Zhu

SAMVIT
Smita Hastak
Charles Yaghmour

NCATS
Raju Hemadri
Nancy Nurthen
Sai Manjula

Adeptia
Sandeep Naredla

Teams
Phenotype & Acquisition
Emily Pfaff, UNC

ACT
Michele Morris, Pitt
Shyam Visweswaran, Pitt
Shawn Murphy HRD

OMOP
Kristin Kostka, IQVIA
Karthik Natarajan, Columbia
Clare Blacketer JNJ

PCORI
Kellie Walters, UNC
Robert Bradford, UNC
Marshall Clark, UNC
Adam Lee, UNC
Evan Colmenares, UNC

TriNetX
Matvey Palchuk
Lora Lingrey

Teams
Analytics
Warren Kibbe, Duke
Heidi Sprait, UTMB
Tell Bennett, U of CO
Andrew Williams, Tufts
Joel Saltz, SBU
Janos Hajagos, SBU
Richard Moffitt, SBU
Tahsin Kurc, SBU

Palantir
Nabeel Qureshi
Andrew Girvin
Amin Manna

Synthetic Data
Regenstrief
Peter Embi

MDClone
Daniel Blumenthal
Hovav Dror
Luz Erez
Josh Rubel

Microsoft
Allison T Rodriguez
Kenji Takeda

National
Covid
Cohort
Collaborative

Thank you!

Questions & Answers

Please submit all questions through the question box.

Summary & Evaluation

- Thank you to all our panelists.
- A recording of the webinar will be available on the CMSS website in the coming weeks.
- Please compete a short evaluation following the webinar.
- For more information, contact info@cmss.org.

CMSS

CMSS WEBINAR SERIES

Advancing Clinical Registries to Support Pandemic Treatment and Response

The series will address key questions related to the rapid development, deployment and implementation of Covid-19 focused clinical registries and clinical repositories by specialty societies and academia.

SUMMER 2020 | FREE TO ATTEND

Upcoming Webinars:

July 17	Reflecting on Our Covid-19 Failures - A New Vision for Integrated Registries
Aug. 6	Deploying Cloud-based Platforms and Analytic Tools to Support Covid-19 and Beyond
Week of Aug. 10	Prioritizing Patient Engagement and Inclusion of Patient-generated Data

CMSS WEBINAR SERIES

Advancing Clinical Registries to Support Pandemic Treatment and Response

Upcoming Webinar:

Reflecting on Our Covid-19 Failures – A New Vision for Integrated Registries

July 17 | 1:30 - 3:00 pm ET

Moderator:

Helen Burstin, MD, MPH, MACP
Chief Executive Officer
Council of Medical Specialty Societies (CMSS)

Panelists:

Elizabeth Garrett-Mayer, PhD
Division Director, Biostatistics and Research Data Governance; Center for Research and Analytics (CENTRA), American Society of Clinical Oncology

Michael Howell, MD, MPH
Principal Scientist, Google
Invited

Clifford Ko, MD, MS, MSHS, FACS, FASCRS
Director, Division of Research and Optimal Patient Care, American College of Surgeons; Vice Chair and Professor of Surgery and Health Services, University of California, Los Angeles (UCLA)

Greg Martin, MD, MSc
Professor of Medicine, Emory University; School of Medicine, Executive Associate Division Director, Division of Pulmonary, Allergy, Critical Care, and Sleep Medicine
President-Elect, Society for Critical Care Medicine