



# **COVID-19 Clinical Data Analysis on a National Scale: *The National COVID Cohort Collaborative (N3C)***

July 29, 2020

*Science through Synergy*

# National COVID Cohort Collaborative (N3C)

The N3C is a partnership among the NCATS-supported [Clinical and Translational Science Awards \(CTSA\) Program](#) hubs and the [National Center for Data to Health \(CD2H\)](#), with overall stewardship by NCATS. Collaborators will contribute and use COVID-19 clinical data to answer critical research questions to address the pandemic.



## **Building an Innovative Analytics Platform to Study COVID-19**

The N3C is a new effort that aims to build a centralized national data resource that the research community can use to study COVID-19 and identify potential treatments as the pandemic continues to evolve. ▶

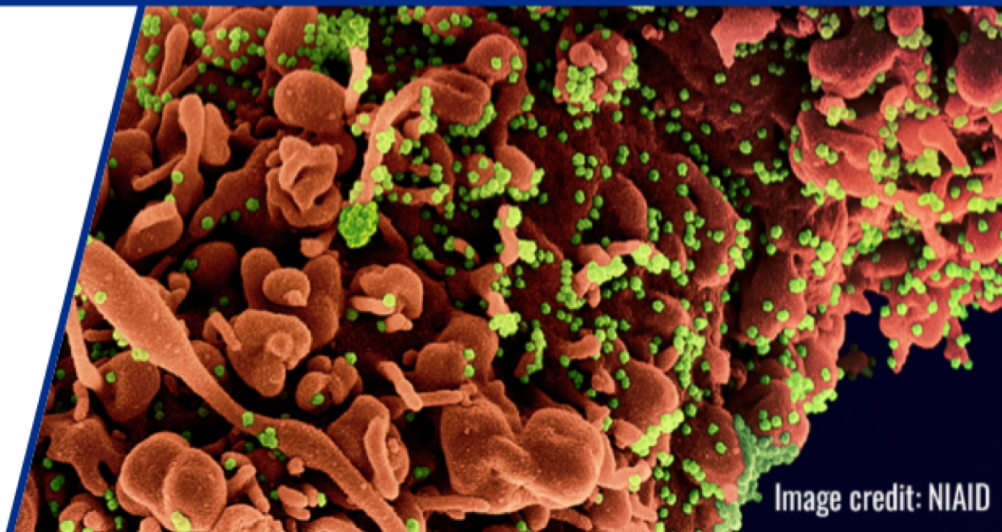
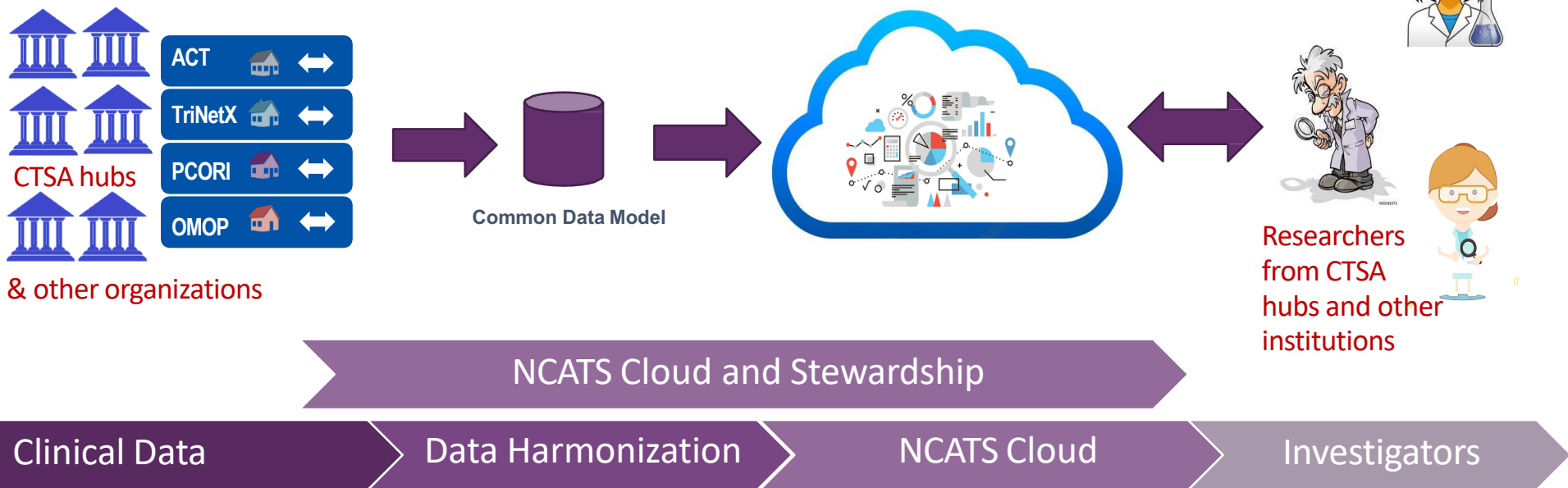


Image credit: NIAID

# How is NCATS doing this?

## Turning Data into Knowledge





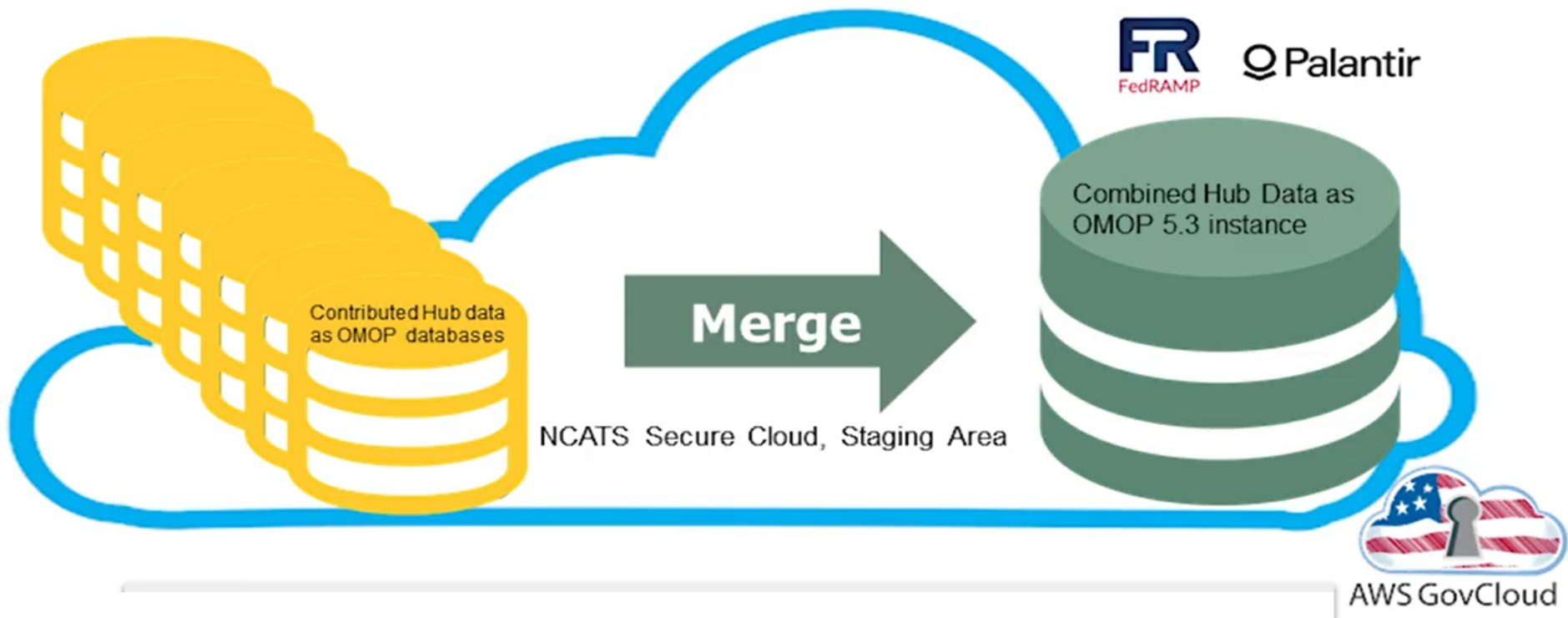
# Electronic Health Record Data



- Person (demographics, location)
- Observation Period
- Specimen
- Death (including cause)
- Visit Occurrence (admissions, visits, services)
- Procedure Occurrence (CPT, ICD10, HCPCS)
- Drug Exposure (orders and administration)
- Device Exposure (instruments, supplies, implantables)
- Condition Occurrence (problems, billing codes)
- Measurement (labs, vital signs, other quantities)
- Observation (other facts such as medical and family history)



# Harmonized to OMOP and Palantir





# N3C Participation



- EHR data contribution
  - Institutional **data transfer** agreement
  - Data transmission: OMOP, i2b2/ACT, PCORNet, TriNetX
  - Ingestions and harmonization
- Analysis of synthetic data set
- Analyze limited data sets
  - Institutional **data use** agreement with NCATS
  - Individual researchers obtain institutional approval (similar to dbGAP)
  - Multi-institution collaborations possible





# Analytics



- Get an N3C Enclave account (join the Analytics Workstream)
- Get project approval
- Analyze limited data sets in the Enclave
- No direct visualization or download of patient-level dataa



# Recent Statistics



COVID-19 Positive  
Patients

16158

Total Patients

309237

Sites Signed DTA

43

Sites Data Ingested

7

Rows of Data

473.5m

Procedures

41.1m

Lab Results

253.6m

Visits

13.3m

Observations

25.1m

Drug Exposures

94.9m



Amin,  
welcome to  
the N3C  
platform!

The N3C Analytical Platform is a collaborative environment for NIH and researchers from the USA's leading institutions to research the novel coronavirus (SARS-CoV-2).

Welcome! To get started:

- 1) For self guided tutorials, [click here](#), or click the question mark on the bottom-left of your screen and click "Academy"!
- 2) Sign up for a training [here](#) and access your training materials [here](#).
- 3) To go through documentation or detailed tutorials for any part of the platform, check out the [documentation here](#).



# Palentir Demonstration

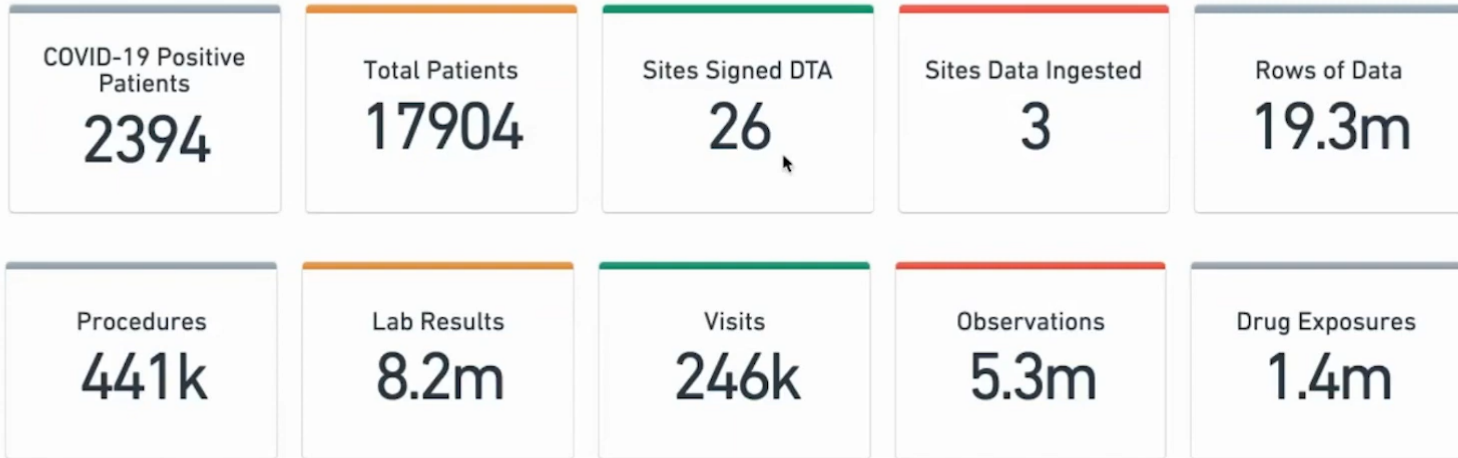
Selected tab

Show Filters

*Christopher G. Chute, MD, DrPH, Johns Hopkins University*

## Key Statistics

N3C Project Statistics / Metrics so far



## Cohort characteristics

Summary statistics for all patients

	COVID (N=2394)	Non-COVID (N=15510)	Overall (N=17904)
<b>Gender</b>			
Male	1122	7178	8299

To plot: Age Group ▾



# Getting Started with N3C



- Have your institution establish a data use agreement
- Institutional data transfer agreement (and data transfer) optional
- Identify collaborators
- Formulate clinical questions
- “On-board” with N3C (<https://cd2h.org/onboard>)
- Obtain IRB exemption for limited data set use
- Obtain institutional approval of your project
- Obtain N3C approval for your project
- Do science





# Data Partnership & Governance Workstream

## Data Access



DAC



Log into  
N3C website



Create an account



Read DUA, UCoC  
CGP, APP



Complete trainings  
IT + Human protection



Describe project



List Collaborators



IRB approval (for LDS)



Sign & submit



Home  
Projects & files  
Recent  
Search... Ctrl + J  
APPS · [View all](#)  
Object explorer  
Contour  
Reports  
Forms  
Slate ☆  
PROJECTS & FILES · [View all](#)  
Data Access  
COVID-19 NIH IRP Invent...  
Notifications  
Help & support  
Account

Hello Mariam,  
please fill out a  
data use  
request!

Create a New Project

Request to become a  
collaborator on an existing  
project

### New Data Use Request

*Each Data Use Request (DUR) will be reviewed by the N3C Data Access Committee (DAC) on a project-specific basis. Requesting Users must submit a DUR for each different proposed Research Project.*

Username and Accessing Institution will be made public.

User Mariam Deacy

\* Accessing Institution

National Center for Advancing Translational Sciences

\* Project Title

Please enter your new project title.

\* Allow other researchers to join this project?

Allow  Do not allow

\* Non-confidential Research Statement ⓘ

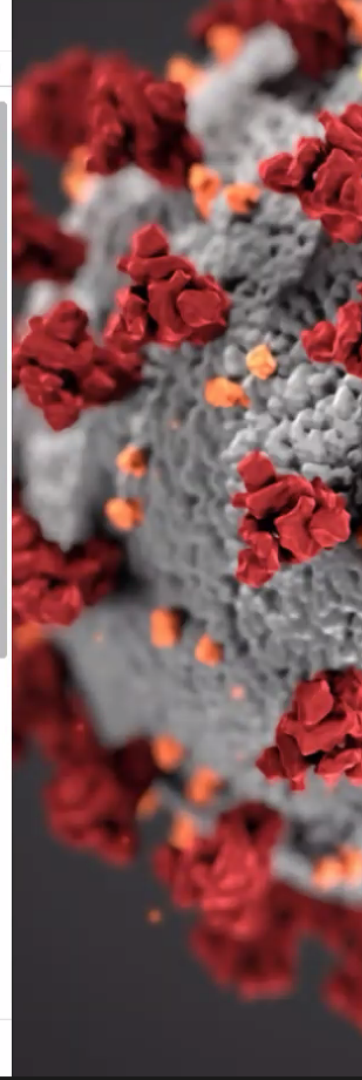
Please type your new non-confidential research statement.

### Collaborators

*Name all project personnel and their institutional affiliations. Note that collaborating Users from different Accessing institutions will need to confirm execution of a DUA between their respective institution and NCATS. Please note that N3C User names and institutional affiliations will be made publicly available on the N3C website.*

3 issues identified

Submit



Home  
Projects & files  
Recent  
Search...  
APPS · View all  
Object explorer  
Contour  
Reports  
Forms  
Slate  
PROJECTS & FILES · View all  
Data Access  
COVID-19 NIH IRP Invent...  
Notifications  
Help & support  
Account

Hello Mariam,  
please fill out a  
data use  
request!

Create a New Project

Request to become a  
collaborator on an existing  
project

### Data Use Request

#### New Data Use Request

Please type your new non-confidential research statement.

#### Collaborators

*Name all project personnel and their institutional affiliations. Note that collaborating Users from different Accessing Institutions will need to confirm execution of a DUA between their respective institution and NCATS. Please note that N3C User names and institutional affiliations will be made publicly available on the N3C website.*

Collaborator Name  
Ken Gersing  
Please enter last name, first name

This collaborator belongs to my institution

Collaborating institution  
Duke University, acting for and on behalf of its School of Medicine

Collaborator Email

+ Add

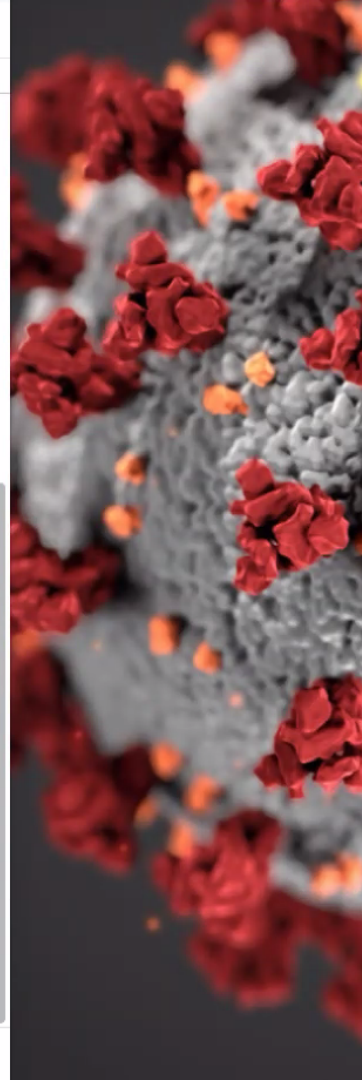
#### Data Tier Access

\* What level of data are you requesting? ⓘ

Limited Data Set (Level 3)  De-Identified Data (Level 2)  Synthetic Data (Level 1)

1 issue identified

Submit





## What variables will be in the enclave?



- That is a great question.
- Will have a better idea as the soft launch of the enclave begins.
- We can make an educated guess based upon what NCATS/N3C requested to be uploaded.
- We are gaining insight by listening N3C weekly workgroup calls.





## What was uploaded?



- [https://ncats.nih.gov/files/OMOP\\_CDM\\_v5\\_COVID\\_042720.pdf](https://ncats.nih.gov/files/OMOP_CDM_v5_COVID_042720.pdf)
- Person demographics (date of birth, gender, race, ethnicity)
- Location (county, state, zip)
- Observations (subject history including smoking status, alcohol)
- Death (date of death)
- Complete Visit History (begin and end dates, Inpatient/Outpatient)



## What was uploaded?



- [https://ncats.nih.gov/files/OMOP\\_CDM\\_v5\\_COVID\\_042720.pdf](https://ncats.nih.gov/files/OMOP_CDM_v5_COVID_042720.pdf)
- History of Prescriptions (begin and end dates, doses, drug)
- History of Device Exposure (X-rays, CT Scans, etc)
- Condition Occurrences (All ICD9, ICD10 codes)
- Measurement Histories (A1C, Glucose, Blood Pressure, Weight, BMI, Creatinine, etc)





## Listening to the calls, what do we hear?



- Data cleaning and quality control procedures.
- Transforming and mapping to constructs (History of Acute Kidney Injury, Type II Diabetes, etc.)
- Calculating comorbidity scores (Charlson, Elixhauser)
- Adding Social Determinants of Health (zip code level)
- Natural Language Processing of the data.
- Data Use Requests – you must propose an analysis and be approved to access the data (synthetic, de-identified, or limited data sets)







# Will you be ready for D3C?



- Test your hypotheses with local data!



<https://www.uab.edu/medicine/scor>



https://www.uab.edu/medicine/scor

# SCOR

SCIENTIFIC COMMUNITY OF OUTCOMES RESEARCHERS

You are here: Home

Providing a home to health services, outcomes, and effectiveness investigators conducting mission-focused research

By cultivating a community of co-located like-minded scholars, structured and unstructured interactions facilitate the transfer of knowledge, promotion of synergies, and generation of novel questions, ideas and collaborations to advance science, wellness, and health, spanning prevention and treatment and diverse content areas. Research training and professional development are central tenets of the SCOR, including dynamic interactions among inter-professional pre- and post-doctoral trainees, early career investigators, senior scientists and staff.

- DEPARTMENT OF MEDICINE
- CENTER FOR CLINICAL AND TRANSLATIONAL SCIENCE
- COERE
- CORE: COLLABORATIVE OUTCOMES RESEARCH ENTERPRISE**
- COVID-19 ENTERPRISE RESEARCH INITIATIVE

**COVID-19 CORE: Collaborative Outcomes Research Enterprise**

The COVID-19 Collaborative Outcomes Research Enterprise (CORE) is leading the initial and longitudinal data collection and information curation to support a broad set of research goals.

Learn More

Collaborative outcomes research enterprise

## COVID-19 CORE

Collaborative Outcomes Research Enterprise

Fueled by the COERE, SCOR & CCTS

- Home
- Organization
- Programmatic Groups
- Scientific Groups

Concept Proposals & Data Requests Resources Questions?

**Purpose:**

To engage a collaborative workforce to coordinate, collate, and facilitate the rigorous and expeditious completion of COVID-19 health services, outcomes and population health studies conducted by interdisciplinary research teams addressing prescient questions germane to the individuals, communities and populations we serve in our city, county, state, and Deep South region.

**Functions:**

- Collation / Clearinghouse of COVID-19 databases (e.g., health system registries, ADPH surveillance)
- Determination and development of specific COVID-19 cohorts / phenotypes (e.g., testing, ICU)
- Concept Proposal coordination: submission, review, feedback (web based, REDCap platform)
- Identify interdisciplinary investigators to form research teams for "approved" Concept Proposals
- Database / Informatics liaison to generate study specific datasets
- BERD liaison to provide feedback and support for design, data analysis, interpretation and dissemination
- Explore multisite collaborative opportunities for Concept Proposals across the CCTS Partner Network

**COERE SCOR CCTS**  
Center for Clinical and Translational Science

Concept proposals & data requests



<https://www.uab.edu/medicine/scor>



UAB THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Scientific Community of Outcomes Researchers  
School of Medicine

Search  Go

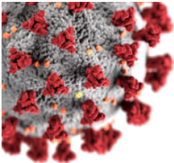
Department Quicklinks ▾ UAB Quicklinks ▾

HOME ABOUT PROGRAMS CONTACT

Information and messages about UAB's re-entry to campus and Ways UAB can unite for our common good

**PROGRAMS**

- LEAD
- GRIT
- DRIVEN
- CORE**



# COVID-19 CORE

Collaborative Outcomes Research Enterprise

Fueled by the COERE, SCOR & CCTS




[Home](#) [Organization](#) [Programmatic Groups](#) [Scientific Groups](#)

[Concept Proposals & Data Requests](#) [Resources](#) [Questions?](#)

[Process & FAQs](#)

[Submit a Concept Proposal](#)

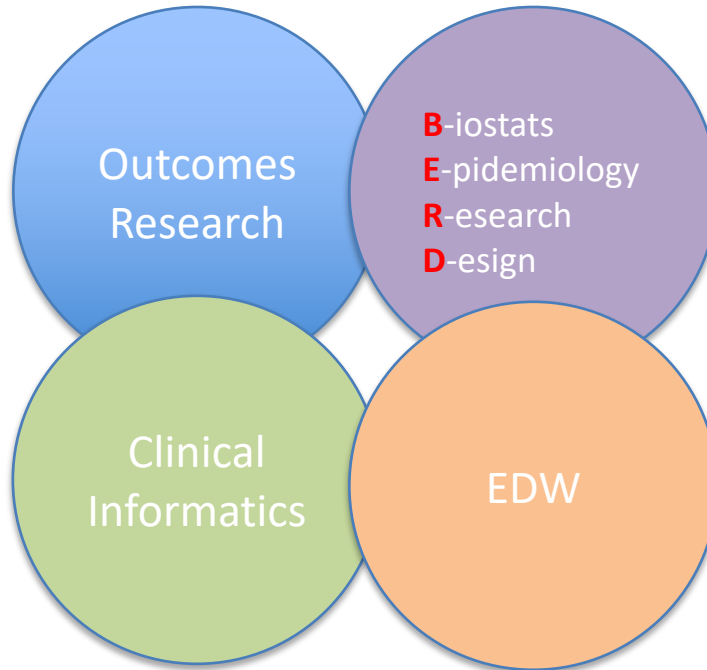
[Request Data Extraction/Transformation](#)

Center for Clinical and Translational Science



# Step 1: Scientific Gateway



- Interdisciplinary feedback
- Emphasis on study design
- Emphasis on analytic plan
- Get to skeleton tables!



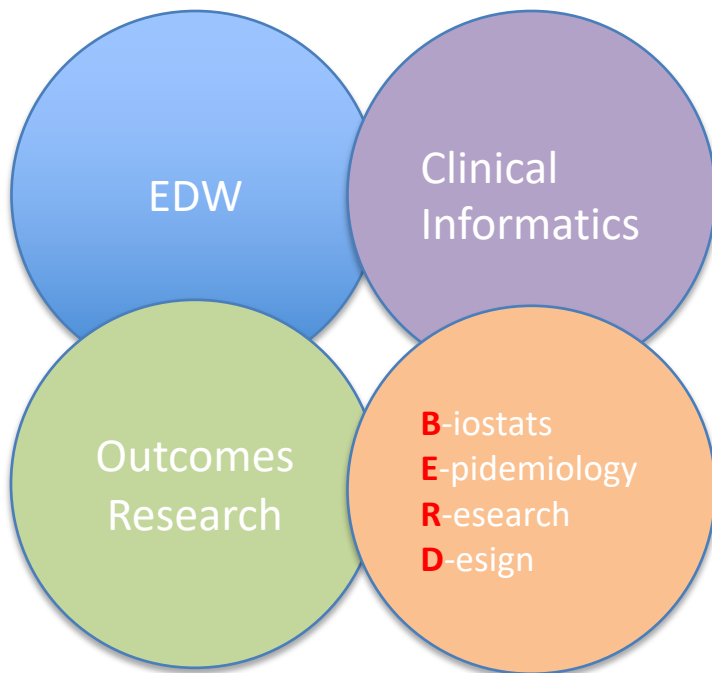
# Scientific Gateway: Concept Proposal



- Investigators; contact info
- Research question
- Background
- Study sample/population
- Outcome, control variables
- Statistical analysis
- Limitations
- Target journal
- References



## Step 2: Data Extraction/Transformation Gateway



- Variables precisely defined
- Data surrogates explored
- Extraction plan
- Transformation plan
- Groupings needed/Plan



# Scientific Gateway: Concept Proposal



- Investigators; contact info
- Research question/objectives
- Purpose (grant, feasibility, etc.)
- Define population
- Outcome, control variables
- Study period (time tether)
- Format for data delivery
  - .CSV, .xlsx
- Mechanism for secure data delivery
- Copy of your IRB







# Extraction/Transformation



- Extraction
  - All patients with a Covid19 RT-PCR Positive test starting 3/1/20-Present
  - All encounters in the EDW for those patients 3/1/20-Present
- Transformation
  - Study population equals: Date of first hospital admission closest in days to first Covid19 RT-PCR
    - Use mappings for inpatient – bedded ER, observation, inpatient
    - Use mappings for all Covid19 RT-PCR tests
  - PT ID/MRN/Date of Positive RT-PCR/Date of first inpatient admission
  - Groupings for Cancer, CHF, DM, etc.





# Next Steps



- Encourage your institution to sign DUA
- Formulate a question
- Identify collaborators
- IRB approval
- Resources:
  - N3C Web site: <https://ncats.nih.gov/n3c>
  - On-boarding: <https://cd2h.org/onboard>

