

**Heidi Spratt PhD, UTMB  
BERD SIG Immediate Past Chair  
hespratt@utmb.edu**

**Heidi Spratt PhD, UTMB  
BERD SIG Immediate Past Chair  
hespratt@utmb.edu**



<https://covid.cd2h.org/>



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# National COVID Cohort Collaborative



NATIONAL CENTER  
FOR DATA TO HEALTH

**The National COVID Cohort Collaborative (N3C) is a complementary and synergistic partnership with overall stewardship by the National Center for Advancing Translational Sciences (NCATS).**



NATIONAL CENTER  
FOR DATA TO HEALTH

**CTSA** Clinical & Translational  
Science Awards Program



National Center  
for Advancing  
Translational Sciences



**ACT**  
The ACT Network



**OHDSI**  
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS



**pcor-net**  
The National Patient-Centered Clinical Research Network



**TriNetX**



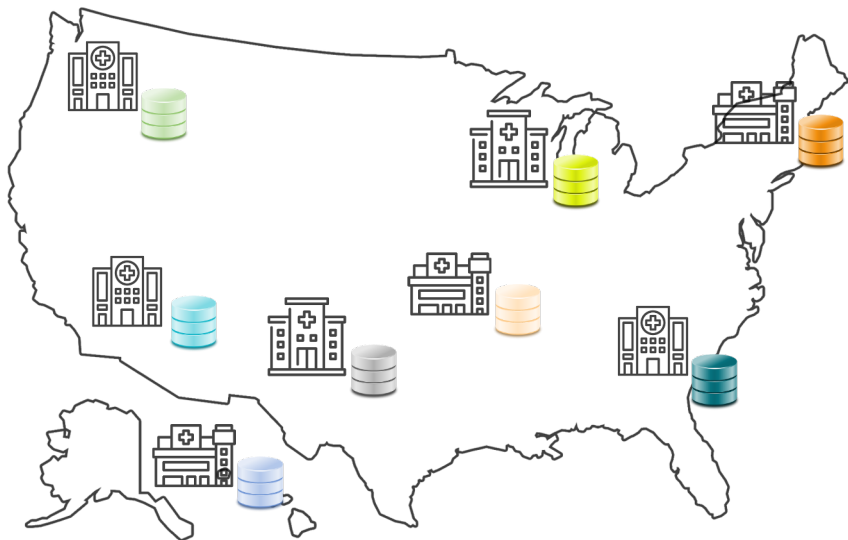
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# National COVID Cohort Collaborative



NATIONAL CENTER  
FOR DATA TO HEALTH





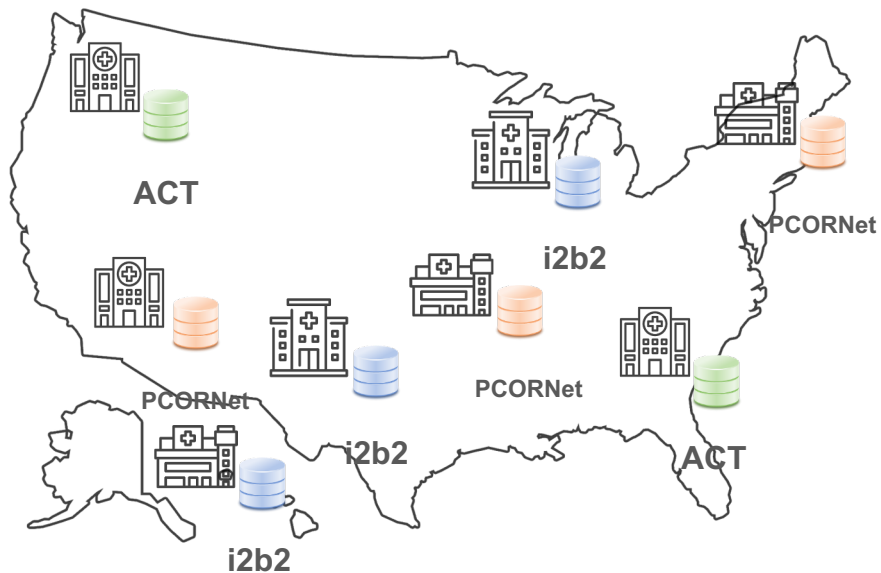
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# National COVID Cohort Collaborative



NATIONAL CENTER  
FOR DATA TO HEALTH







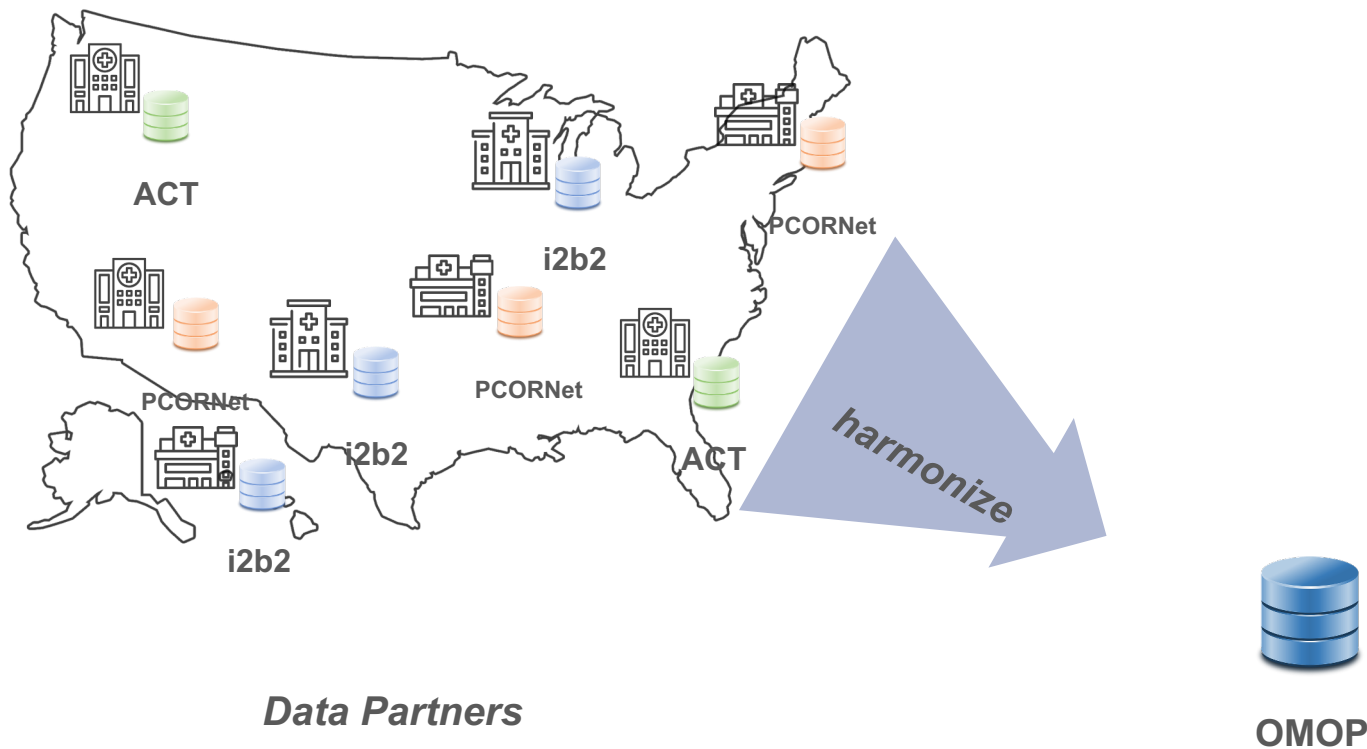
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# National COVID Cohort Collaborative



NATIONAL CENTER  
FOR DATA TO HEALTH





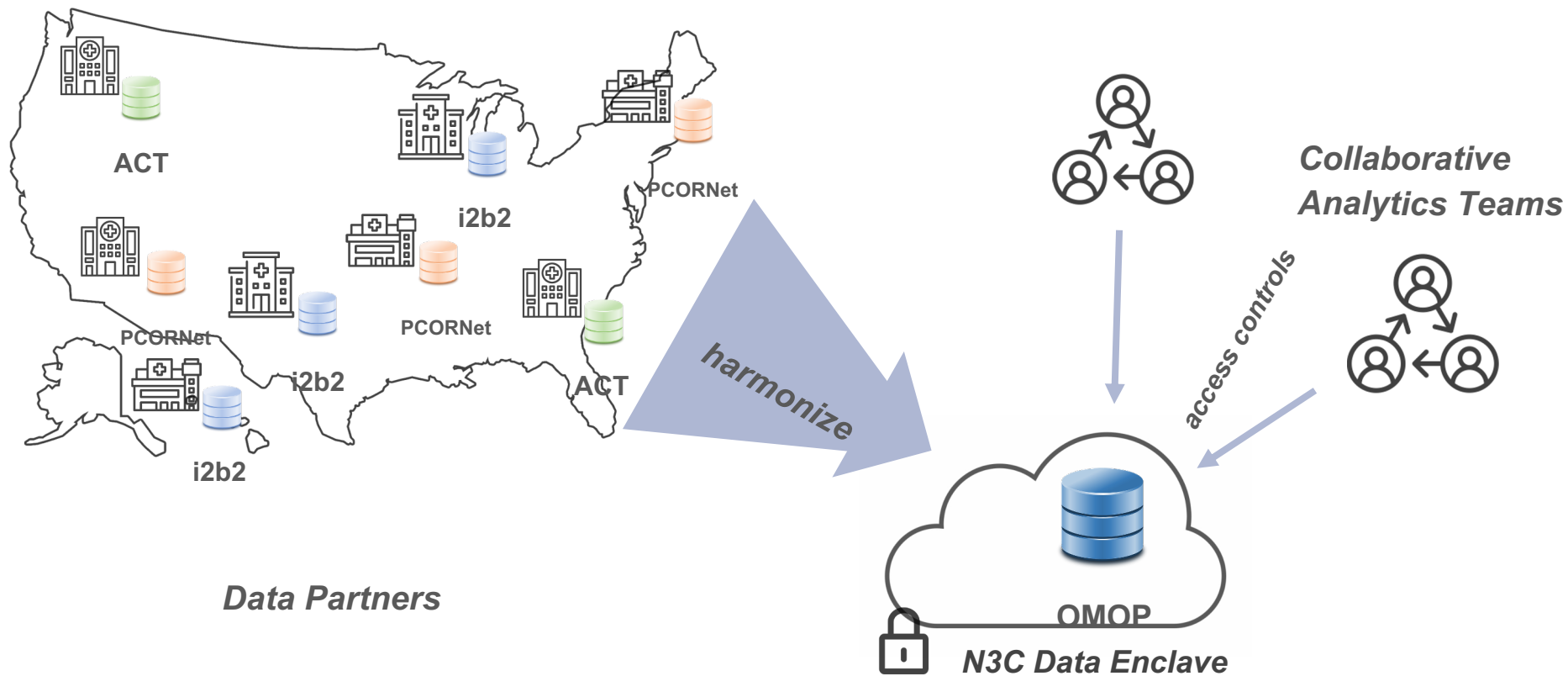
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# National COVID Cohort Collaborative



NATIONAL CENTER  
FOR DATA TO HEALTH





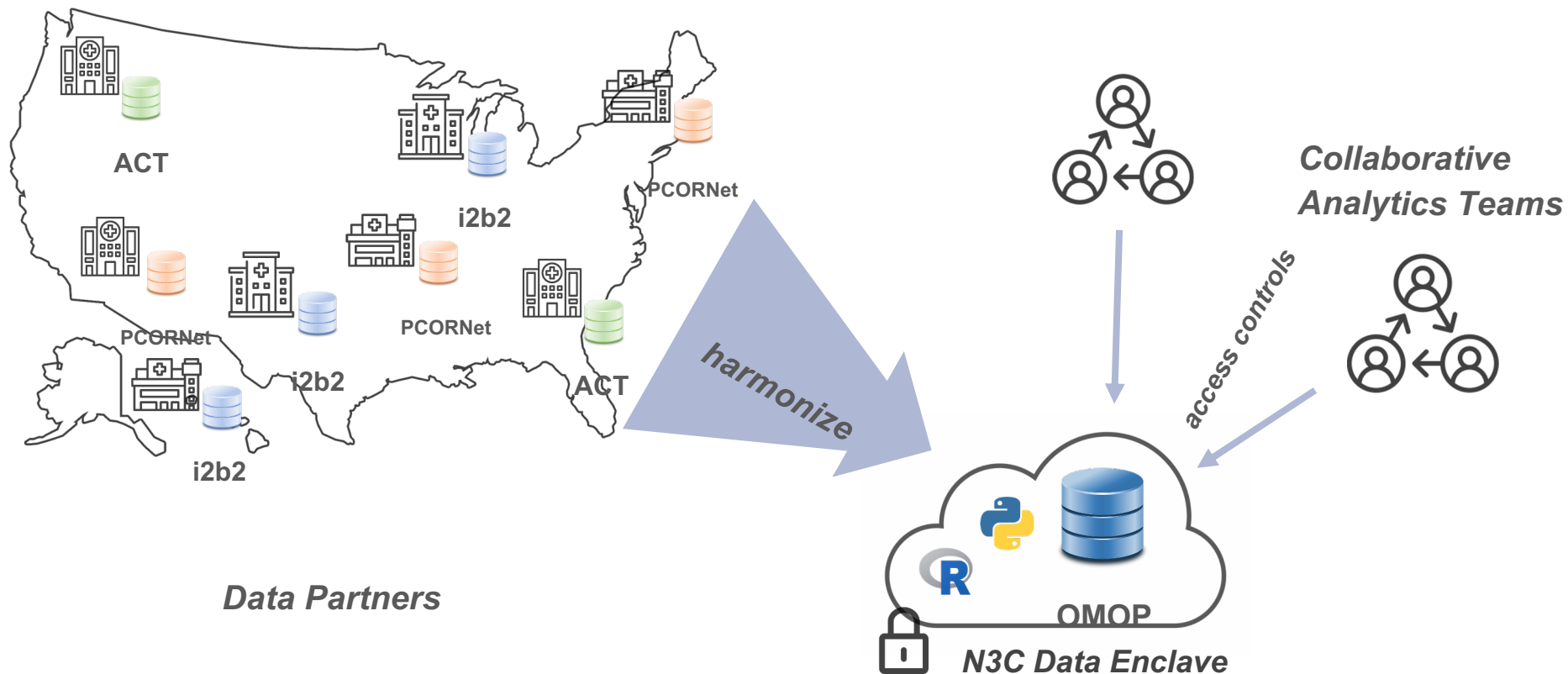
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# National COVID Cohort Collaborative



NATIONAL CENTER  
FOR DATA TO HEALTH





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C is an Enterprise



NATIONAL CENTER  
FOR DATA TO HEALTH

**Research  
organizations  
and individuals**

**2,000+  
participants  
across 500+  
institutions**

**FedRAMP cert, SSO,  
auditing, tool  
deployment, workflow  
management, hashing**

**Technical  
infrastructure**

**Workstreams,  
Subgroups,  
Domain teams,  
G-suite, Slack,  
calendaring,  
listservs,  
website,  
newsletters**

**Service desk,  
ticket  
system,  
training  
materials &  
events,  
analytics,  
visualization,  
HPC**

**Governance &  
regulatory**

**Legal agreements, IRBs,  
DTAs, DUAs, DURs, DAC,  
community & public-  
private partnerships**

**EHR Data, Data  
Harmonization, QC,  
CDM tools,  
Codesets, variables**

**Data management &  
Knowledge  
engineering**





# What kinds of questions can N3C address?



*The scope and scale of the information in the platform will support probing questions such as:*

- *Do some therapies work **better than** others? By region? By demographics?*
- *Can we **compare** local rare clinical observations with national occurrences?*
- *Can we **predict** who might have severe outcomes if they have COVID-19?*
- *Can we **predicate** acute renal failure in COVID+ patients?*
- ***What** factors will predict the effectiveness of vaccines?*
- ***Who** might need a ventilator **because of** lung failure?*
- ***Why** are some people susceptible to Long COVID?*
- *For kids, can we **compare** rates of infectivity before/after school reopening?*
- *What social determinants of health are **risk factors** for mortality?*



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# National COVID Cohort Collaborative



NATIONAL CENTER  
FOR DATA TO HEALTH

 **1,222,296<sup>\*</sup>**

POSITIVE PATIENTS

 **50**

SITES

 **4,960,128**

TOTAL PATIENTS

 **5.8b**

TOTAL ROWS

**287.3m**

PROCEDURES

**949.0m**

DRUG EXPOSURES

**257.8m**

VISITS

**721.4m**

OBSERVATIONS

**2.6b**

LAB RESULTS

COVID+ patients and 2:1 matched “controls” – matched based on age, race, gender, and ethnicity

All records 1/1/2018 and later

\* Count of positive COVID-19 cases is defined as the number of distinct patients who have at least one positive result for any COVID-19 lab test (diagnostic or antibody) or a diagnosis of U07.1, the ICD-10-CM code for *COVID-19, virus identified*.



National  
COVID  
Cohort  
Collaborative

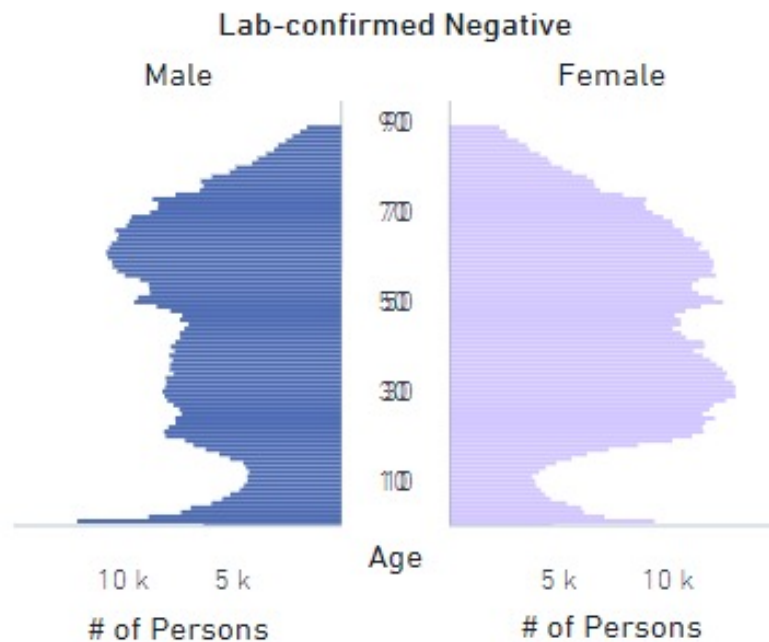
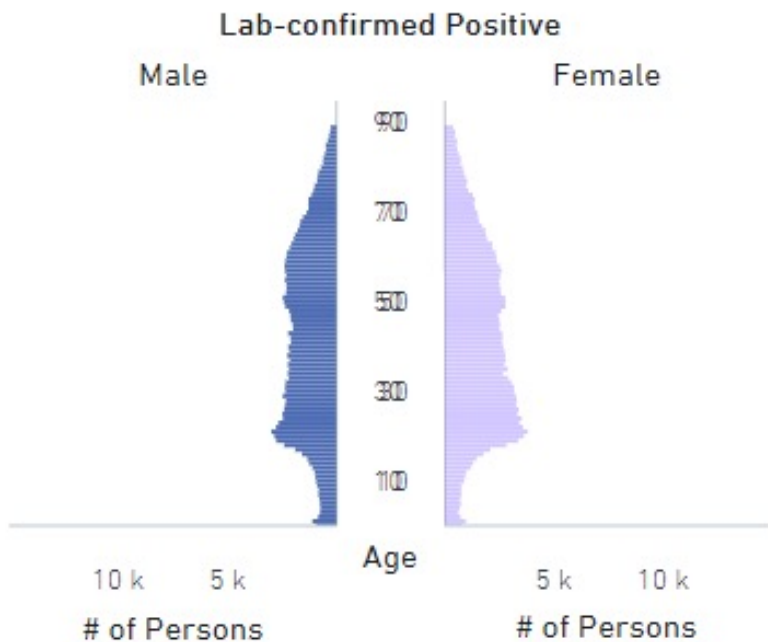
A program of NIH's National Center  
for Advancing Translational Sciences

# Dashboard



NATIONAL CENTER  
FOR DATA TO HEALTH

## Cohort by age and gender





# Dashboard



<https://covid.cd2h.org/dashboard/>

## Comorbidity Distribution of COVID+ in N3C Cohort

Condition	Mild			Hospitalized			All COVID+	Hospitalized
	Mild	Mild ED	Moderate	Severe	Mortality/Hospice			
Diabetes Mellitus	9.00%	12.6%	27.4%	37.6%	39.5%		12.9%	29.3%
Renal Disease	3.07%	4.30%	15.6%	19.8%	32.0%		5.71%	17.7%
Congestive Heart Failure	2.05%	3.29%	13.1%	17.8%	28.8%		4.44%	15.2%
Chronic Pulmonary Disease	8.91%	12.7%	19.2%	20.4%	26.7%		11.3%	20.1%
Peripheral Vascular Disease	3.62%	5.54%	13.1%	12.4%	24.1%		5.67%	14.3%
Stroke	2.19%	3.32%	9.84%	11.8%	20.1%		3.85%	11.1%
Cancer	3.67%	3.37%	9.37%	7.76%	17.5%		4.75%	10.2%
Dementia	0.550%	0.820%	4.95%	3.51%	16.2%		1.54%	6.14%
Heart Attack	1.14%	2.12%	6.86%	9.95%	15.0%		2.42%	7.97%
Liver Disease	2.62%	3.72%	7.74%	10.1%	11.7%		3.72%	8.33%
Rheumatologic Disease	2.43%	2.80%	4.69%	3.97%	5.87%		2.88%	4.78%
Partial Paralysis/Weakness	0.290%	0.460%	2.01%	3.54%	3.77%		0.650%	2.30%
Totals	(n = 200,890)	(n = 32,509)	(n = 39,828)	(n = 2,795)	(n = 5,489)		(n = 281,511)	(n = 48,112)





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Data Access Levels



NATIONAL CENTER  
FOR DATA TO HEALTH

*Level 3*



**OMOP**

Data Level	Data Description
Limited Data Set (LDS)	Patient data that retain the following protected health information — <ul style="list-style-type: none"><li>• dates of service</li><li>• patient ZIP code</li></ul>



# Data Access Levels



**Level 3**



**OMOP**

Data Level	Data Description
Limited Data Set (LDS)	Patient data that retain the following protected health information — <ul style="list-style-type: none"><li>• dates of service</li><li>• patient ZIP code</li></ul>
De-identified Data Set	Patient data from the LDS with the following changes — <ul style="list-style-type: none"><li>• Dates of service are algorithmically shifted to protect patient privacy.</li><li>• Patient ZIP codes are truncated to the first three digits or removed entirely if the ZIP code represents fewer than 20,000 individuals.</li></ul>

**Level 2**



**OMOP**



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Data Access Levels



NATIONAL CENTER  
FOR DATA TO HEALTH

Level 3



OMOP

Data Level	Data Description
Limited Data Set (LDS)	Patient data that retain the following protected health information — <ul style="list-style-type: none"><li>• dates of service</li><li>• patient ZIP code</li></ul>
De-identified Data Set	Patient data from the LDS with the following changes — <ul style="list-style-type: none"><li>• Dates of service are algorithmically shifted to protect patient privacy.</li><li>• Patient ZIP codes are truncated to the first three digits or removed entirely if the ZIP code represents fewer than 20,000 individuals.</li></ul>
Synthetic Data Set	Data that are computationally derived from the LDS that resemble patient information statistically but are not actual patient data.

Level 2



OMOP

Level 1



Custom



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Data Access Levels



NATIONAL CENTER  
FOR DATA TO HEALTH

Level 3



OMOP

Data Level	Data Description	Eligible Users
Limited Data Set (LDS)	Patient data that retain the following protected health information — <ul style="list-style-type: none"><li>• dates of service</li><li>• patient ZIP code</li></ul>	<ul style="list-style-type: none"><li>• Researchers from U.S.-based institutions</li></ul>
De-identified Data Set	Patient data from the LDS with the following changes — <ul style="list-style-type: none"><li>• Dates of service are algorithmically shifted to protect patient privacy.</li><li>• Patient ZIP codes are truncated to the first three digits or removed entirely if the ZIP code represents fewer than 20,000 individuals.</li></ul>	<ul style="list-style-type: none"><li>• Researchers from U.S.-based institutions</li><li>• Researchers from foreign institutions</li></ul>
Synthetic Data Set	Data that are computationally derived from the LDS that resemble patient information statistically but are not actual patient data.	<ul style="list-style-type: none"><li>• Researchers from U.S.-based institutions</li><li>• Researchers from foreign institutions</li><li>• Citizen scientists</li></ul>

Level 2



OMOP

Level 1



Custom





# Data Access Levels

Level 3



OMOP

Level 2



OMOP

Level 1



Custom

Data Level	Data Description	Eligible Users
Limited Data Set (LDS)	<p>Patient data that retain the following protected health information —</p> <ul style="list-style-type: none"> <li>• dates of service</li> <li>• patient ZIP code</li> </ul>	<ul style="list-style-type: none"> <li>• Researchers from U.S.-based institutions</li> </ul>
De-identified Data Set	<p>Patient data from the LDS with the following changes —</p> <ul style="list-style-type: none"> <li>• Dates of service are algorithmically shifted to protect patient privacy.</li> <li>• Patient ZIP codes are truncated to the first three digits or removed entirely if the ZIP code represents fewer than 20,000 individuals.</li> </ul>	<ul style="list-style-type: none"> <li>• Researchers from U.S.-based institutions</li> <li>• Researchers from foreign institutions</li> </ul>
Synthetic Data Set	<p>Data that are computationally derived from the LDS that resemble patient information statistically but are not actual patient data.</p>	<ul style="list-style-type: none"> <li>• Researchers from U.S.-based institutions</li> <li>• Researchers from foreign institutions</li> <li>• Citizen scientists</li> </ul>

- N3C registration (1)
- N3C Data Enclave account (2)
- DUA executed with NCATS (for 2)
- NIH IT training completion (for 3)
- Approved DUR (3)



# Data Access Levels

Level 3



OMOP

Data Level	Data Description	Eligible Users
Limited Data Set (LDS)	Patient data that retain the following protected health information — <ul style="list-style-type: none"> <li>• dates of service</li> <li>• patient ZIP code</li> </ul>	<ul style="list-style-type: none"> <li>• Researchers from U.S.-based institutions</li> </ul>

Level 2



OMOP

De-identified Data Set	Patient data from the LDS with the following changes — <ul style="list-style-type: none"> <li>• Dates of service are algorithmically shifted to protect patient privacy.</li> <li>• Patient ZIP codes are truncated to the first three digits or removed entirely if the ZIP code represents fewer than 20,000 individuals.</li> </ul>	<ul style="list-style-type: none"> <li>• Researchers from U.S.-based institutions</li> <li>• Researchers from foreign institutions</li> </ul>	<ul style="list-style-type: none"> <li>• N3C registration</li> <li>• N3C Data Enclave account</li> <li>• DUA executed with NCATS</li> <li>• NIH IT training completion</li> <li>• Approved DUR</li> <li>• Human Subjects Research Protection training completion</li> </ul> <p>(Have completion date for DUR)</p>
------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Level 1



Custom

Synthetic Data Set	Data that are computationally derived from the LDS that resemble patient information statistically but are not actual patient data.	<ul style="list-style-type: none"> <li>• Researchers from U.S.-based institutions</li> <li>• Researchers from foreign institutions</li> <li>• Citizen scientists</li> </ul>	<ul style="list-style-type: none"> <li>• N3C registration (1)</li> <li>• N3C Data Enclave account (2)</li> <li>• DUA executed with NCATS (for 2)</li> <li>• NIH IT training completion (for 3)</li> <li>• Approved DUR (3)</li> </ul>
--------------------	-------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



# Data Access Levels

Level 3



OMOP

Level 2



OMOP

Level 1



Custom

Data Level	Data Description	Eligible Users	Access Requirements*
Limited Data Set (LDS)	Patient data that retain the following protected health information — <ul style="list-style-type: none"> <li>• dates of service</li> <li>• patient ZIP code</li> </ul>	<ul style="list-style-type: none"> <li>• Researchers from U.S.-based institutions</li> </ul>	<ul style="list-style-type: none"> <li>• N3C registration</li> <li>• N3C Data Enclave account</li> <li>• Data Use Agreement (DUA) executed with NCATS</li> <li>• NIH IT training completion</li> <li>• Approved Data Use Request (DUR)</li> <li>• Human Subjects Research Protection training completion</li> <li>• Local Human Research Protection Program IRB determination letter</li> </ul>
De-identified Data Set	Patient data from the LDS with the following changes — <ul style="list-style-type: none"> <li>• Dates of service are algorithmically shifted to protect patient privacy.</li> <li>• Patient ZIP codes are truncated to the first three digits or removed entirely if the ZIP code represents fewer than 20,000 individuals.</li> </ul>	<ul style="list-style-type: none"> <li>• Researchers from U.S.-based institutions</li> <li>• Researchers from foreign institutions</li> </ul>	<ul style="list-style-type: none"> <li>• N3C registration</li> <li>• N3C Data Enclave account</li> <li>• DUA executed with NCATS</li> <li>• NIH IT training completion</li> <li>• Approved DUR</li> <li>• Human Subjects Research Protection training completion</li> </ul> <p>(Have completion date for DUR)</p>
Synthetic Data Set	Data that are computationally derived from the LDS that resemble patient information statistically but are not actual patient data.	<ul style="list-style-type: none"> <li>• Researchers from U.S.-based institutions</li> <li>• Researchers from foreign institutions</li> <li>• Citizen scientists</li> </ul>	<ul style="list-style-type: none"> <li>• N3C registration (1)</li> <li>• N3C Data Enclave account (2)</li> <li>• DUA executed with NCATS (for 2)</li> <li>• NIH IT training completion (for 3)</li> <li>• Approved DUR (3)</li> </ul>



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Data Usage Agreement (DUA)



NATIONAL CENTER  
FOR DATA TO HEALTH

- **Agreement to behave!**

[https://ncats.nih.gov/files/NCATS\\_N3C\\_Data\\_Use\\_Agreement.pdf](https://ncats.nih.gov/files/NCATS_N3C_Data_Use_Agreement.pdf)

- COVID-research only
- don't re-identify patients or share data
- everything in the enclave is logged
- & more



National Center  
for Advancing  
Translational Sciences

<https://covid.cd2h.org/account-instructions>





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Data Usage Agreement (DUA)



NATIONAL CENTER  
FOR DATA TO HEALTH

- **Agreement to behave!**

[https://ncats.nih.gov/files/NCATS\\_N3C\\_Data\\_Use\\_Agreement.pdf](https://ncats.nih.gov/files/NCATS_N3C_Data_Use_Agreement.pdf)

- COVID-research only
- don't re-identify patients or share data
- everything in the enclave is logged
- & more



National Center  
for Advancing  
Translational Sciences

- **Many institutions sign on behalf of their employees**

- 150+ so far: <https://covid.cd2h.org/duas>
- The Ohio State University has a signed DUA

- **Individuals can sign as well**

- Citizen scientists - directly with NIH NCATS

<https://covid.cd2h.org/account-instructions>



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Enclave & Data Access



NATIONAL CENTER  
FOR DATA TO HEALTH

## N3C Data Enclave (unite.nih.gov)

Enclave Access:  
Registration + DUA  
coverage confirmed



Level 3  
(LDS)



Level 2  
(De-IDd)



Level 1  
(Synthetic)



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Enclave & Data Access



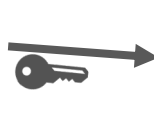
NATIONAL CENTER  
FOR DATA TO HEALTH

Enclave Access:  
Registration + DUA  
coverage confirmed

## N3C Data Enclave (unite.nih.gov)



Research  
Project



Level 3  
(LDS)

Research  
Project



Level 2  
(De-IDd)

Research  
Project



Level 1  
(Synthetic)



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

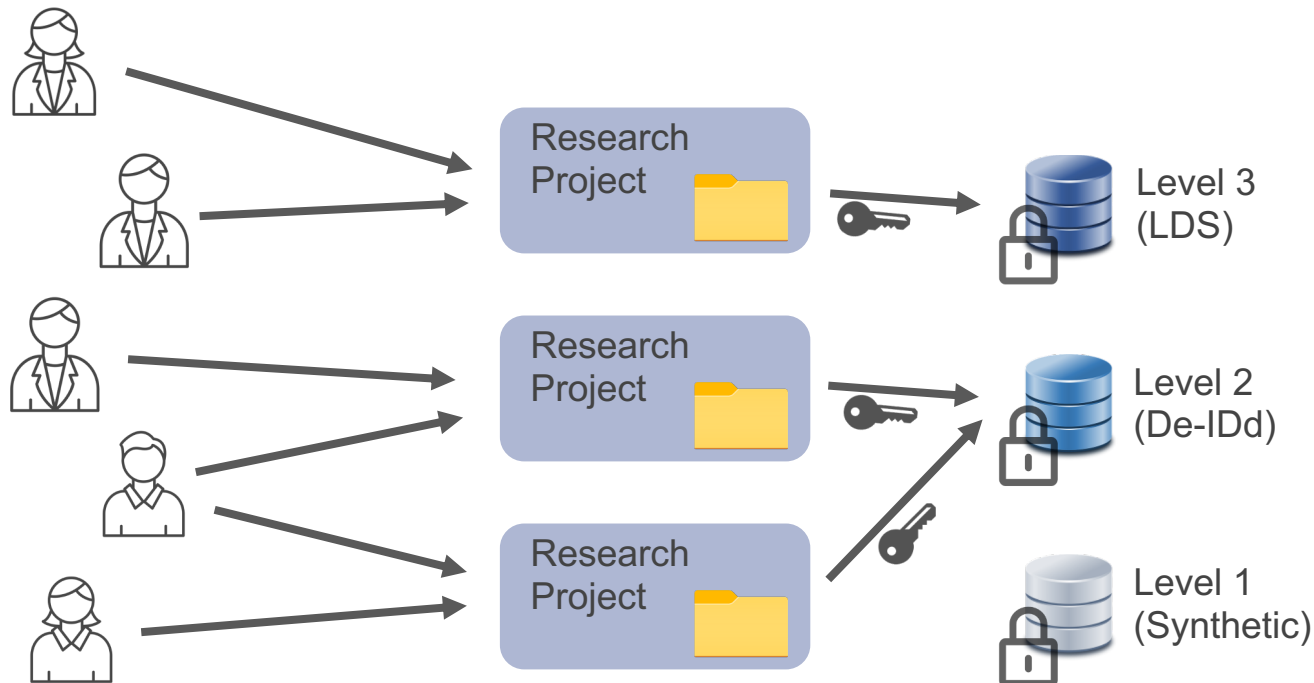
# Enclave & Data Access



NATIONAL CENTER  
FOR DATA TO HEALTH

Enclave Access:  
Registration + DUA  
coverage confirmed

## N3C Data Enclave (unite.nih.gov)





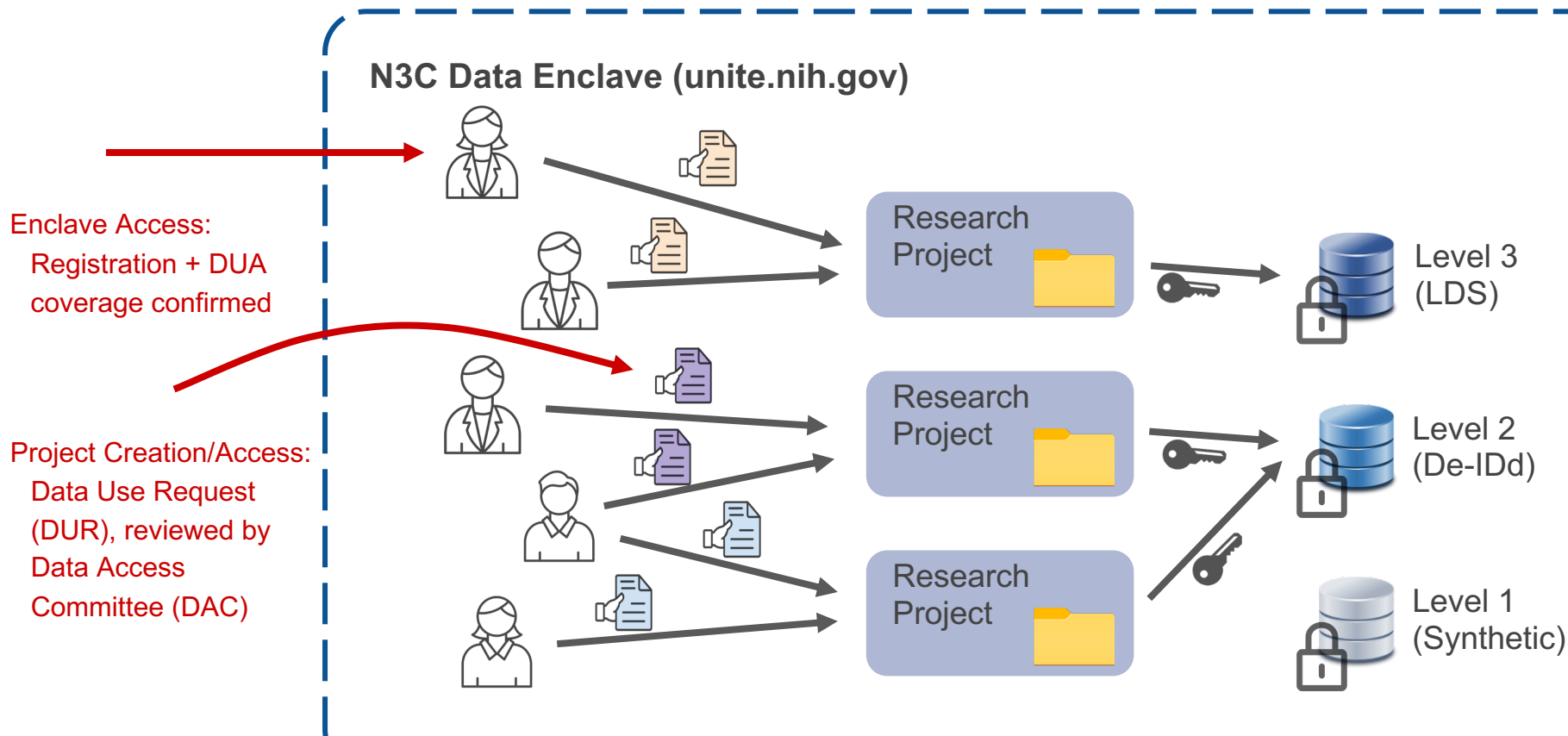
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Enclave & Data Access



NATIONAL CENTER  
FOR DATA TO HEALTH





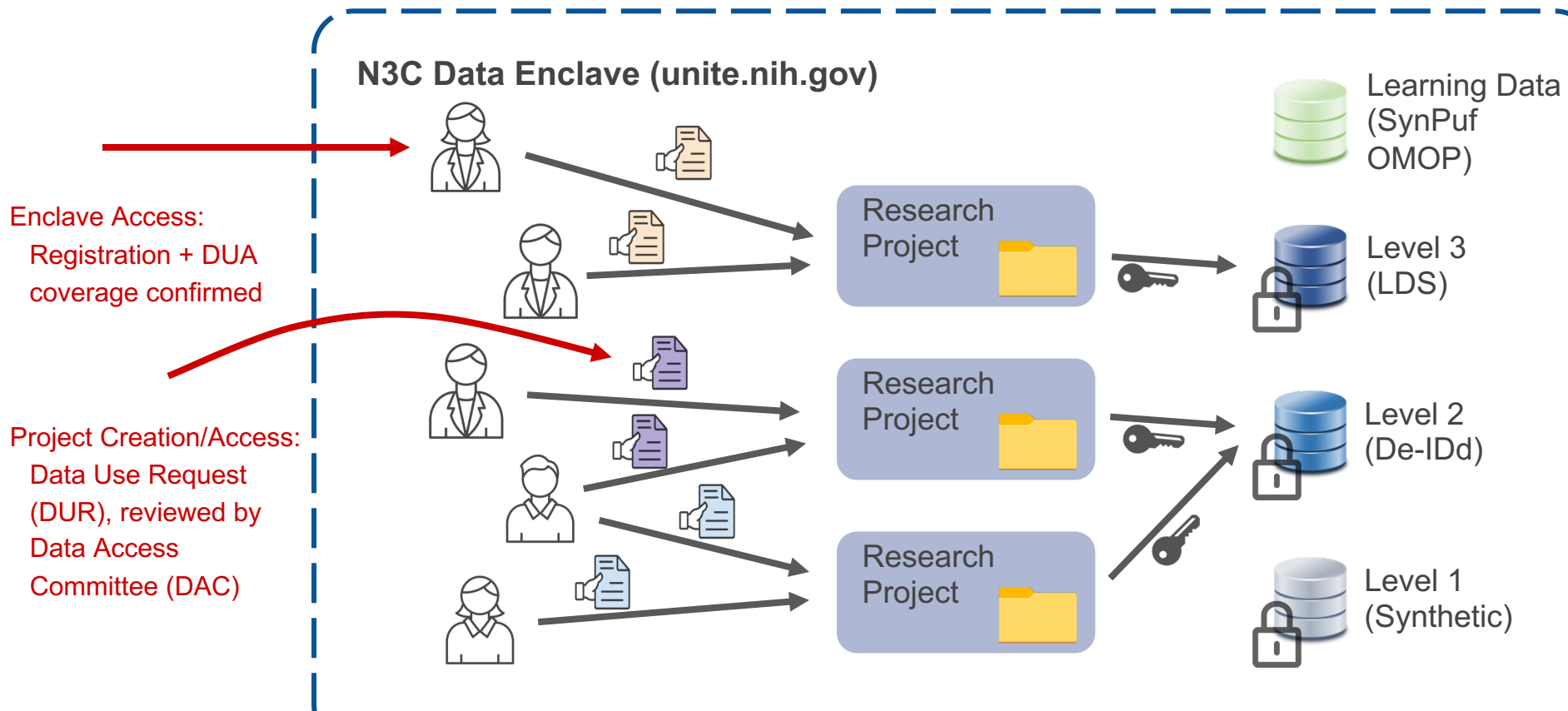
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Enclave & Data Access



NATIONAL CENTER  
FOR DATA TO HEALTH





# Other Rules of the Road



- **User Code of Conduct:** <https://ncats.nih.gov/n3c/resources/data-user-code-of-conduct>
  - COVID-research only
  - Follow laws and regulations
  - Don't re-identify people, institutions, or populations
  - No screenshots of data or results, no recording video calls that show data
    - Video calls w/ data shown are OK if all on the call are approved & have access themselves
  - Do publish, do attribute N3C properly
  - Report data access incidents within 2 business days





# Other Rules of the Road



- **User Code of Conduct:** <https://ncats.nih.gov/n3c/resources/data-user-code-of-conduct>
  - COVID-research only
  - Follow laws and regulations
  - Don't re-identify people, institutions, or populations
  - No screenshots of data or results, no recording video calls that show data
    - Video calls w/ data shown are OK if all on the call are approved & have access themselves
  - Do publish, do attribute N3C properly
  - Report data access incidents within 2 business days
- **Community Guiding Principles:** <https://doi.org/10.5281/zenodo.3978728>
  - *Partnership, Inclusivity, Transparency, Reciprocity, Accountability, Security*
  - Also guidelines on Ethics, Diversity and Inclusion, and Resolving Issues through Community Arbitration



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Other Rules of the Road



NATIONAL CENTER  
FOR DATA TO HEALTH

## ● Publications & Exporting Results

- Got a paper to *submit*? Poster or abstract *accepted*? Get it reviewed by the Publication Committee!
- All summary tables and figures must be submitted (via form in the enclave) for review before export  
Summary: no “cell sizes” of 1 through 19 (0 and 20+ are OK)
- Details at <https://covid.cd2h.org/publication-review>



# Other Rules of the Road





## ● Publications & Exporting Results


- Got a paper to *submit*? Poster or abstract *accepted*? Get it reviewed by the Publication Committee!
- All summary tables and figures must be submitted (via form in the enclave) for review before export  
Summary: no “cell sizes” of 1 through 19 (0 and 20+ are OK)
- Details at <https://covid.cd2h.org/publication-review>


## ● External/Supplemental Data (e.g. census or other public data)


- External data can be used alongside N3C level 1/2/3 data, *but*:
- It must be reviewed for security and licence and uploaded by admins
- All previously uploaded external data are made available to all (must acknowledge risk to access)
- Details at <https://covid.cd2h.org/external-datasets>


# N3C Data Enclave




 Home


 Projects & files


 Recent


 Search... Ctrl + J


★ APPS · [View all](#)


 Object explorer


 Contour


 Reports


 Code workbook

 Code repositories


 Data lineage


 Issues


 Builds


 Slate ☆

★ PROJECTS & FILES · [View all](#)

 flights

 Data Quality Portal

 Notifications

 Help & support

HS Account



National COVID Cohort Collaborative Data Enclave  
Discover, collaborate, and analyze data from the N3C.



NIH National Center  
for Advancing  
Translational Sciences

## Welcome to N3C, Heidi!

1,222,296  
POSITIVE PATIENTS

50  
SITES

4,960,128  
TOTAL PATIENTS

5.8b  
TOTAL ROWS

287.3m  
PROCEDURES

949.0m  
DRUG EXPOSURES

257.8m  
VISITS

721.4m  
OBSERVATIONS

2.6b  
LAB RESULTS

### Quick links

Data use request (DUR)

My projects

Explore projects

Support & Contact

Training material

Data access overview

### Applications

Phenotype explorer

Concept set browser

Get Help!  
Also see

[https://covid.cd2h.org/help\\_desk](https://covid.cd2h.org/help_desk)






National  
COVID  
Cohort  
Collaborative





NATIONAL CENTER  
FOR DATA TO HEALTH


# N3C Data Enclave




 Home


 Projects & files


 Recent


 Search... Ctrl + J


★ APPS · [View all](#)


 Object explorer


 Contour


 Reports


 Code workbook

 Code repositories


 Data lineage


 Issues


 Builds


 Slate ☆

★ PROJECTS & FILES · [View all](#)

 flights

 Data Quality Portal

 Notifications

 Help & support

HS Account



National COVID Cohort Collaborative Data Enclave  
Discover, collaborate, and analyze data from the N3C.



NIH National Center  
for Advancing  
Translational Sciences

Welcome to N3C, Heidi!

1,222,296  
POSITIVE PATIENTS

50  
SITES

4,960,128  
TOTAL PATIENTS

5.8b  
TOTAL ROWS

287.3m  
PROCEDURES

949.0m  
DRUG EXPOSURES

257.8m  
VISITS

721.4m  
OBSERVATIONS

2.6b  
LAB RESULTS

## Quick links

Data use request (DUR)

My projects

Explore projects

Support & Contact

Training material

Data access overview

## Applications

Phenotype explorer

Concept set browser

Training  
Portal, etc.

Get Help!  
Also see

[https://covid.cd2h.org/help\\_desk](https://covid.cd2h.org/help_desk)

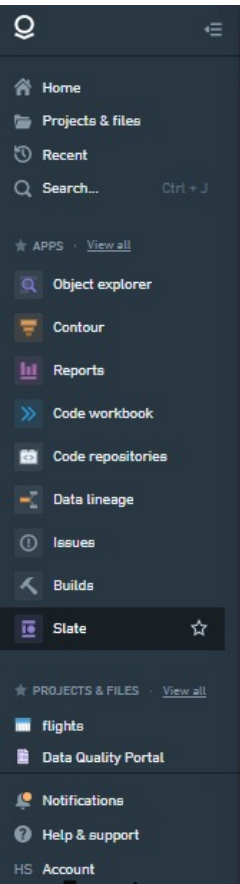


National  
COVID  
Cohort  
Collaborative



NATIONAL CENTER  
FOR DATA TO HEALTH

# N3C Data Enclave



- Home
- Projects & files
- Recent
- Search... Ctrl + J
- APPS · [View all](#)
- Object explorer
- Contour
- Reports
- Code workbook
- Code repositories
- Data lineage
- Issues
- Builds
- Slate
- PROJECTS & FILES · [View all](#)
- flights
- Data Quality Portal
- Notifications
- Help & support
- HS Account



National COVID Cohort Collaborative Data Enclave  
Discover, collaborate, and analyze data from the N3C.



NIH National Center  
for Advancing  
Translational Sciences

Welcome to N3C, Heidi!

1,222,296

POSITIVE PATIENTS

50

SITES

4,960,128

TOTAL PATIENTS

5.8b

TOTAL ROWS

287.3m

PROCEDURES

949.0m

DRUG EXPOSURES

257.8m

VISITS

721.4m

OBSERVATIONS

2.6b

LAB RESULTS

Quick links



Data use request (DUR)



My projects



Explore projects



Support & Contact



Training materials



Data access overview

Applications



Phenotype explorer



Concept set browser

More details:

<https://covid.cd2h.org/tutorials>  
(esp. "Enclave Organization")

DUR Management

Training  
Portal, etc.

Get Help!  
Also see

[https://covid.cd2h.org/help\\_desk](https://covid.cd2h.org/help_desk)



National  
COVID  
Cohort  
Collaborative



NATIONAL CENTER  
FOR DATA TO HEALTH

# N3C Data Enclave

The screenshot shows the N3C Data Enclave dashboard. On the left is a dark sidebar with navigation links: Home, Projects & files, Recent, Search..., APPS (with a 'View all' link), Object explorer, Contour, Reports, Code workbook, Code repositories, Data lineage, Issues, Builds, Slate, PROJECTS & FILES (with a 'View all' link), flights, Data Quality Portal, Notifications, Help & support, and HS Account. The main content area has a header with the N3C logo and text: 'National COVID Cohort Collaborative Data Enclave' and 'Discover, collaborate, and analyze data from the N3C.' The top right shows the NIH logo and 'National Center for Advancing Translational Sciences'. The main content area starts with a 'Welcome to N3C, Heidi!' message. Below this is a row of five statistics: 1,222,296 POSITIVE PATIENTS, 50 SITES, 4,960,128 TOTAL PATIENTS, 5.8b TOTAL ROWS, and 287.3m PROCEDURES. Below this is another row of five statistics: 949.0m DRUG EXPOSURES, 257.8m VISITS, 721.4m OBSERVATIONS, and 2.6b LAB RESULTS. Below the statistics is a 'Quick links' section with six buttons: Data use request (DUR), My projects, Explore projects, Support & Contact, Training material, and Data access overview. Below the quick links is an 'Applications' section with two cards: Phenotype explorer and Concept set browser. Red arrows point from text annotations to various parts of the dashboard: 'Project folders, files, data, etc.' points to 'Projects & files' in the sidebar; 'Analysis tools' points to 'Object explorer', 'Contour', 'Reports', and 'Code workbook' in the sidebar; 'DUR Management' points to the 'Data use request (DUR)' button in the quick links; 'Training Portal, etc.' points to the 'Training material' button in the quick links; 'Get Help! Also see' points to the 'Support & Contact' button in the quick links. A blue text annotation 'More details: https://covid.cd2h.org/tutorials (esp. "Enclave Organization")' points to the 'TOTAL ROWS' statistic. Another blue text annotation 'https://covid.cd2h.org/help\_desk' points to the 'Support & Contact' button. The bottom left corner has the N3C logo and 'National COVID Cohort Collaborative'. The bottom right corner has the NIH logo and 'NATIONAL CENTER FOR DATA TO HEALTH'.

**Project folders, files, data, etc.**

**Analysis tools**

**DUR Management**

**Training Portal, etc.**

**Get Help!**  
Also see [https://covid.cd2h.org/help\\_desk](https://covid.cd2h.org/help_desk)

More details:  
<https://covid.cd2h.org/tutorials>  
(esp. "Enclave Organization")

**Quick links**

- Data use request (DUR)
- My projects
- Explore projects
- Support & Contact
- Training material
- Data access overview

**Applications**

- Phenotype explorer
- Concept set browser

**Statistics**

Statistic	Value
POSITIVE PATIENTS	1,222,296
SITES	50
TOTAL PATIENTS	4,960,128
TOTAL ROWS	5.8b
PROCEDURES	287.3m
DRUG EXPOSURES	949.0m
VISITS	257.8m
OBSERVATIONS	721.4m
LAB RESULTS	2.6b

**Footer**

National COVID Cohort Collaborative

NATIONAL CENTER FOR DATA TO HEALTH





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Screenshots!



NATIONAL CENTER  
FOR DATA TO HEALTH

The screenshot shows a web browser window with the URL `unite.nih.gov/workspace/compass/projects`. The page has a dark sidebar on the left with navigation links: Home, Projects & files, Recent, Search..., APPS (with a View all link), Object explorer, Contour, Reports, Slate, Code workbook, Code repositories, Issues, PROJECTS & FILES (with a View all link), Training Portal V2, and Training Portal. The main content area has tabs for Data Catalog, Projects, and Shared with you. The Projects tab is active, showing a 'Recommended' section with three project cards: '[Testing] Identified Data', 'Safe Harbor', and '[EXTDATASET-4] Social Deprivat...'. Below this is a search bar 'Filter projects...' and a table of projects. The table has columns for NAME, LAST UPDATED, and TAGS. The projects listed are: '[EXTDATASET-2] Area Health Resources Files', '[EXTDATASET-4] Social Deprivation Index (SDI)', '[RP-01A7C4] Retrospective identification of medications to treat COVID-19', and '[RP-020D1B] Analysis of Time Windows as Determining Factors for COVID-19 P...'. Each project card has a 'Request access' button.

NAME	LAST UPDATED	TAGS
[EXTDATASET-2] Area Health Resources Files <a href="https://data.hrsa.gov/topics/health-workforce/ahrf">https://data.hrsa.gov/topics/health-workforce/ahrf</a>		Request access
[EXTDATASET-4] Social Deprivation Index (SDI) [Robert Graham Center] "...developed the Social Deprivation Index (SDI) to quantify L...	Fri, Jan 8, 2021, 12:06:50 PM	
[RP-01A7C4] Retrospective identification of medications to treat COVID-19 COVID-19 is a disease caused by the novel coronavirus SARS-CoV-2 which has infec...		Request access
[RP-020D1B] Analysis of Time Windows as Determining Factors for COVID-19 P... Protocol that determines if a patient should be tested for COVID-19 may depend on ...		Request access

## Project Folders/Workspaces (Navigation)



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Screenshots!



NATIONAL CENTER  
FOR DATA TO HEALTH

The screenshot shows the 'Projects' page in the unite.nih.gov workspace. The left sidebar contains navigation links: Home, Projects & files, Recent, Search..., APPS (with a 'View all' link), Object explorer, Contour, Reports, Slate, Code workbook, Code repositories, and Issues. The main content area has tabs for Data Catalog, Projects, and Shared with you. Under the 'Projects' tab, there's a 'Recommended' section with three project cards: '[Testing] Identified Data', 'Safe Harbor', and '[EXTDATASET-4] Social Deprivat...'. Below this is a search bar 'Filter projects...' and dropdowns for 'Your role' and 'Template'. A table lists projects with columns 'NAME', 'LAST UPDATED', and 'TAGS'. The projects listed are: '[EXTDATASET-2] Area Health Resources Files', '[EXTDATASET-4] Social Deprivation Index (SDI)', '[RP-01A7C4] Retrospective identification of medications to treat COVID-19', and '[RP-020D1B] Analysis of Time Windows as Determining Factors for COVID-19 P...'. Each project has a 'Request access' button. A large red 'X' is drawn over these buttons. To the right of the table, a red text box says 'Ignore these "Request access" buttons'.

NAME	LAST UPDATED	TAGS
[EXTDATASET-2] Area Health Resources Files https://data.hrsa.gov/topics/health-workforce/ahrf		
[EXTDATASET-4] Social Deprivation Index (SDI) [Robert Graham Center] "...developed the Social Deprivation Index (SDI) to quantify L...	Fri, Jan 8, 2021, 12:06:50 PM	
[RP-01A7C4] Retrospective identification of medications to treat COVID-19 COVID-19 is a disease caused by the novel coronavirus SARS-CoV-2 which has infec...		
[RP-020D1B] Analysis of Time Windows as Determining Factors for COVID-19 P... Protocol that determines if a patient should be tested for COVID-19 may depend on ...		

## Project Folders/Workspaces (Navigation)

Data Catalog

unite.nih.gov/workspace/compass/data-catalog

Home  
Projects & files  
Recent  
Search...  
★ APPS · View all  
Object explorer  
Contour  
Reports  
Slate  
Code workbook  
Code repositories  
Issues

Data Catalog

Projects Shared with you

Data Catalog

Collections Files

NAME ^

- ☒ Dashboards and Apps  
Important applications and dashboards
- ☒ De-Identified Data  
Harmonized De-Identified Data
- ☒ External Datasets
- ☒ Foundry Academy Tutorials
- ☒ N3C Knowledge Store  
N3C Shared Logic. All logic is shared and open
- ☒ OMOP Code Sets  
Sets of OMOP Codes
- ☐ OMOP Concepts



Browser: pneumonia\_fluoroquinolones | x +

Address bar: unite.nih.gov/workspace/contour-app/ri.contour.main.analysis.b857e448-3af2-4075-a274-6b42229f2c27/path/ri.contour.main.ref.0692f8c2-a0fd-4...

Page Title: N3C Training ... > pneumonia\_fluoroquinolones

Buttons: File Help, Editing mode, Share

Navigation: pneumonia, fluoroquinolones, joined, visualize, +

Toggle: Show table

Notification: New version of data available. [Details] [Update data]

14	314795728	1712572	fluoroquinolones [...]	***	true	1	trovafloxacin 100 ...	***
15	314795728	1716903	fluoroquinolones [...]	***	true	1	moxifloxacin	***
16	314795728	1716905	fluoroquinolones [...]	***	true	1	moxifloxacin 400 ...	***
17	314795728	1742253	fluoroquinolones [...]	***	true	1	levofloxacin	***
18	314795728	1742255	fluoroquinolones [...]	***	true	1	levofloxacin 500 M	***

TEXT [Add to report] [Edit] [X]

As with the pneumonia path, we inner-join with the data table of interest, this time `drug_era`.

JOIN [Settings] [X]

Intersection with `drug_era` where `concept_id` matches `drug_concept_id`

Calculate summary

GRID [Add to report] [Settings] [X]

Unique count of `person_id` by `exposure_days` and `drug_concept_name`



[[RP-CA3365]] ... > Sandbox > Shawn > test\_analysis ☆

File ▾ Help ▾ | [[RP-CA3365]] N3C Immu... | Output ▾ Branch: master ▾

Environment ▾ Parameters ▾ ⚙ ▾ Actions ▾ 0 12 3 | Share

CONTENTS

Paths Graph

+ New Transform Import Dataset Manual Entry Delete

pan / Select Layout Colors

**Dataset 1**

person_id	gender	concept_id	year_of_birth	month_of_birth	day_of_birth
76	M	8532	1938	3	1
331	M	8532	1945	7	1
353	M	8532	1953	11	1
431	M	8532	1921	9	1
477	M	8532	1937	1	1

person as person2  
25 columns • 116,352 rows

**Dataset 2**

person_id	drug_exposure_id	drug_concept_id	drug_exposure_start_date
4	288	2213440	2009-06-24
4	289	2213440	2009-06-24
5	295	2213440	2009-09-24
7	324	2213440	2010-03-19
7	323	2213440	2009-04-30

drug\_exposure as drug\_e...  
28 columns • 4,294,108 rows

**Dataset 3**

concept_id	concept_name	is_most_recent
7579872	5084 ARIScience - Outpa...	false
80002939	5084 ARIScience - Outpa...	true

concept\_set\_members as  
6 columns • 1,549,811 rows

List of all concept sets used further to obtain concept sets via template for a particular... Read more

**HIV PATIENTS FROM CONDITI...**  
No parameter values

**HIV PATIENTS FROM DRUG E...**  
No parameter values

**HIV PATIENTS COMBIN...**  
No parameter values

**all\_hiv\_patients\_flags**  
15 columns

**result**  
15 columns

Preview all\_hiv\_patients\_flags  
15 columns

Save as dataset Toggle view Actions

Inputs: person2 hiv\_b20\_z21\_patients hiv\_drug\_patients hiv\_lab\_patients hiv\_condition\_patients

```

1 def all_hiv_patients_flags(hiv_condition_patients, hiv_lab_patients, hiv_drug_patients, hiv_b20_z21_patients, person2):
2
3     hiv_condition_patients.createOrReplaceGlobalTempView("hiv_condition_patients")
4     hiv_lab_patients.createOrReplaceGlobalTempView("hiv_lab_patients")
5     hiv_drug_patients.createOrReplaceGlobalTempView("hiv_drug_patients")
6     hiv_b20_z21_patients.createOrReplaceGlobalTempView("hiv_b20_z21_patients")

```

CONSOLE





National  
COVID  
Cohort  
Collaborative

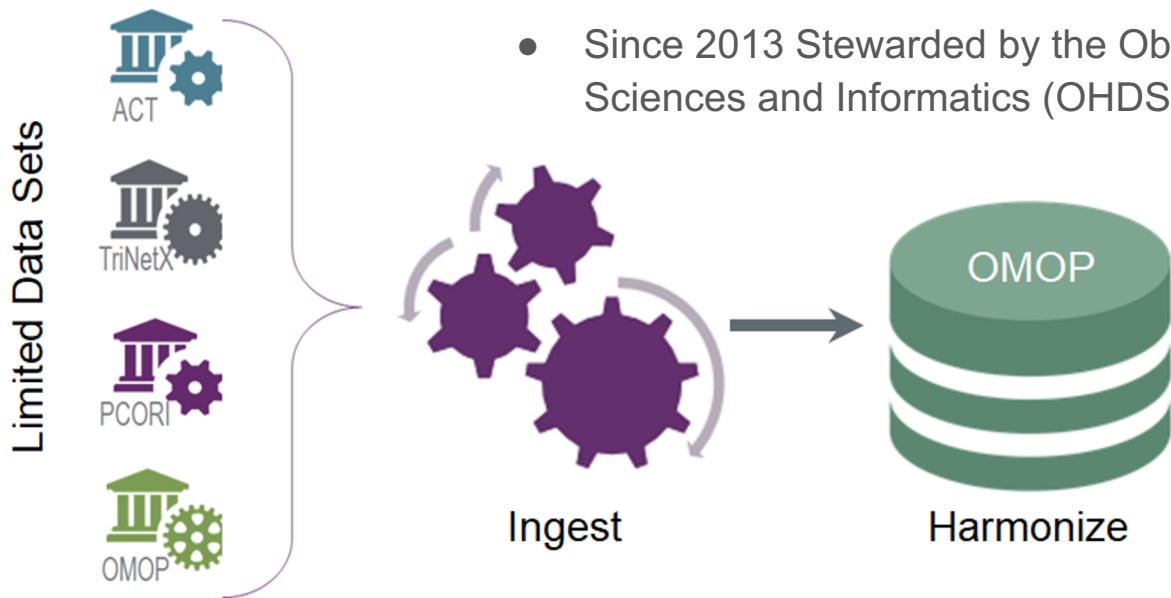
A program of NIH's National Center  
for Advancing Translational Sciences

# Why OMOP?



NATIONAL CENTER  
FOR DATA TO HEALTH

- OMOP - Observational Medical Outcomes Partnership  
(funding expired 2013)
- Since 2013 Stewarded by the Observational Health Data  
Sciences and Informatics (OHDSI) collaborative





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# OMOP Design



NATIONAL CENTER  
FOR DATA TO HEALTH

- **Purpose Driven:** Organize data for *analysis*, not billing or other purposes.
- **Data protection:** Default to excluding identifying information (names, birthdates), with exceptions where warranted.



# OMOP Design



- **Purpose Driven:** Organize data for *analysis*, not billing or other purposes.
- **Data protection:** Default to excluding identifying information (names, birthdates), with exceptions where warranted.
- **Design of domain tables:** The data domains (tables) are *person- and event-centric*; most records include a person ID and date (there are also tables describing providers, person demographics, and locations).





# OMOP Design



- **Purpose Driven:** Organize data for *analysis*, not billing or other purposes.
- **Data protection:** Default to excluding identifying information (names, birthdates), with exceptions where warranted.
- **Design of domain tables:** The data domains (tables) are *person- and event-centric*; most records include a person ID and date (there are also tables describing providers, person demographics, and locations).

drug_exposure								Showing 300 of 6.3m rows	
	person_id Long	drug_exposure_id Long	drug_concept_id Integer	drug_exposure_star... Date	drug_exposure_star... Timestamp	drug_exposure_end_... Date	drug Tim		
1	4	288	2213440	2009-06-24	009-06-24T00:00:00.000Z	2009-06-24			
2	4	289	2213440	2009-06-24	009-06-24T00:00:00.000Z	2009-06-24			
3	5	295	2213440	2009-09-24	009-09-24T00:00:00.000Z	2009-09-24			
4	7	324	2213440	2010-03-19	010-03-19T00:00:00.000Z	2010-03-19			
5	7	323	2213440	2009-04-30	009-04-30T00:00:00.000Z	2009-04-30			



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# OMOP Design



NATIONAL CENTER  
FOR DATA TO HEALTH

- **Reuse of Existing Vocabularies:** OMOP “concepts” are leveraged from national or industry standards or vocabularies. Every concept (e.g. J18.9, the ICD-10 code for pneumonia) is also given an OMOP-specific ID (e.g. 35207956).
- **Provenance & Standards:** Original data values are kept, but *mapped* to a set of standards identified by OMOP for analysis



# OMOP Design



- **Reuse of Existing Vocabularies:** OMOP “concepts” are leveraged from national or industry standards or vocabularies. Every concept (e.g. J18.9, the ICD-10 code for pneumonia) is also given an OMOP-specific ID (e.g. 35207956).
- **Provenance & Standards:** Original data values are kept, but *mapped* to a set of standards identified by OMOP for analysis
- **Standardized Vocabulary Hierarchy:** These preferred standard terms live in a hierarchy for analysis use: “ProQuin XR” Has Ingredient “Ciprofloxacin 500 Mg” Is A “Antibiotic” - *use this hierarchy to define a **concept set** and collect data records of interest!*



# OMOP Design



- **Reuse of Existing Vocabularies:** OMOP “concepts” are leveraged from national or industry standards or vocabularies. Every concept (e.g. J18.9, the ICD-10 code for pneumonia) is also given an OMOP-specific ID (e.g. 35207956).
- **Provenance & Standards:** Original data values are kept, but *mapped* to a set of standards identified by OMOP for analysis
- **Standardized Vocabulary Hierarchy:** These preferred standard terms live in a hierarchy for analysis use: “ProQuin XR” Has Ingredient “Ciprofloxacin 500 Mg” Is A “Antibiotic” - *use this hierarchy to define a **concept set** and collect data records of interest!*
- **OMOP Tooling:** OHDSI and N3C have developed special tools for working with OMOP data.  
[Enclave Training Portal](#) -> [N3C Modules](#) -> [Intro to OMOP & Contour](#)



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Data Harmonization



NATIONAL CENTER  
FOR DATA TO HEALTH

- The OMOP (Observational Medical Outcomes Partnership) model aligns Electronic Health Record data into a standard database format and terminology

## **“Pneumonia”**



J18.9 (ICD10 code)



486 (ICD9 code)



D0110146 (MeSH ID)



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

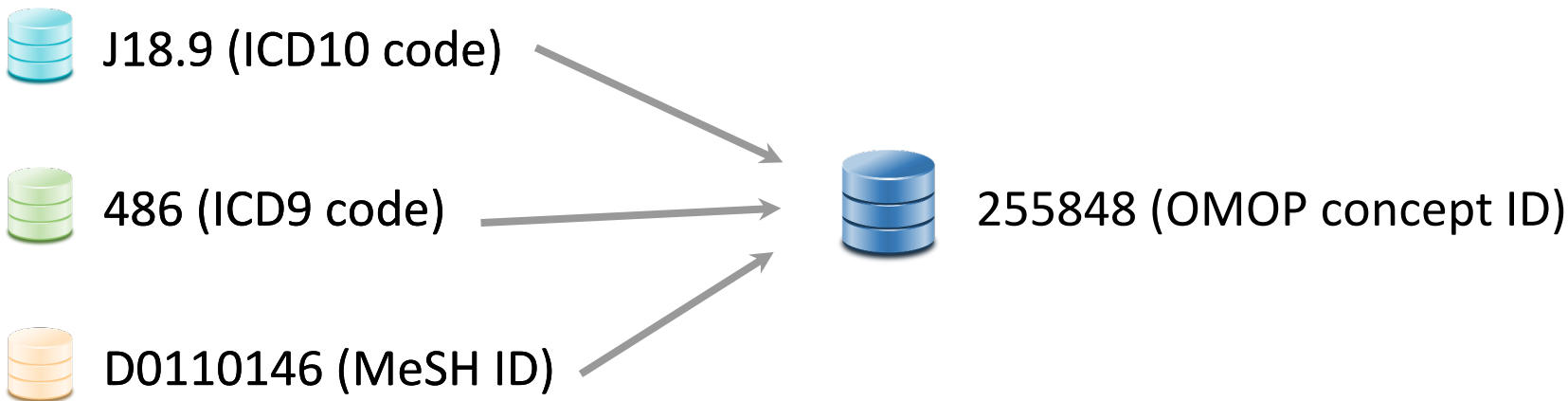
# Data Harmonization



NATIONAL CENTER  
FOR DATA TO HEALTH

- The OMOP (Observational Medical Outcomes Partnership) model aligns Electronic Health Record data into a standard database format and terminology

## “Pneumonia”



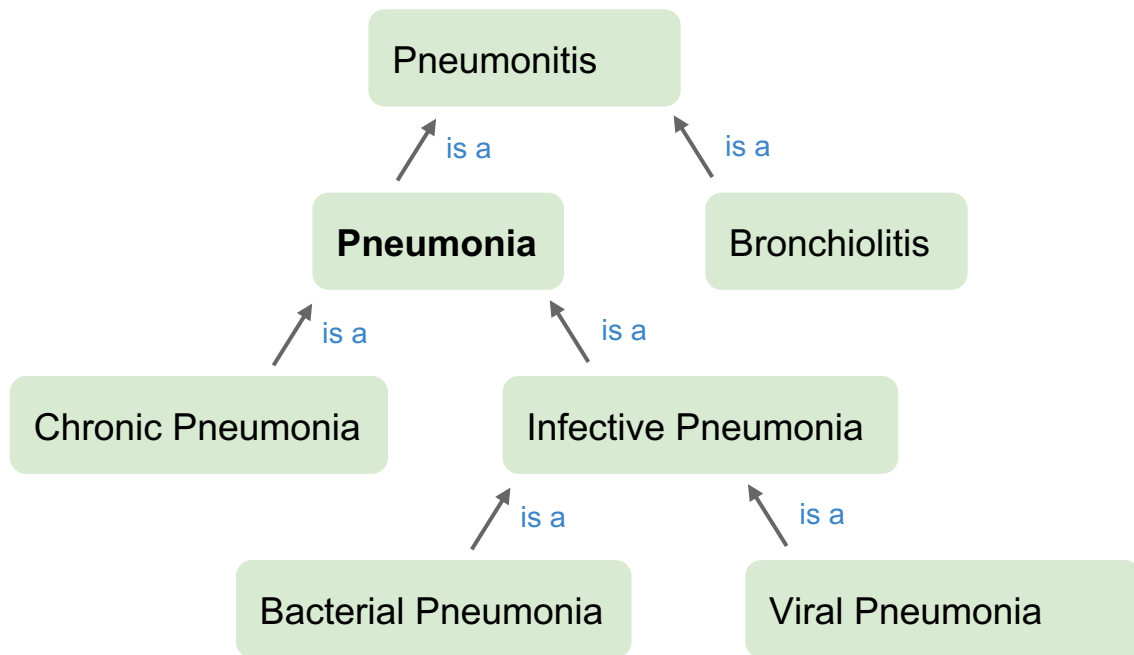
*Note: messy details not shown*



# OMOP Terminology



- Within the standards, there's a hierarchy and relationships



*"...all cases of pneumonia, except subtypes of chronic pneumonia..."*

*"...all exposures to antibiotics containing ciprofloxacin ..."*



*"...all individuals taking cipro within 10 days of a chronic pneumonia diagnosis ..."*



# OMOP/N3C Data Tables



- Person
  - Year of birth
  - Gender
  - Race
  - Ethnicity
  - Location (State + 3- or 5-digit zip code depending on access level)
- Visit
  - Start/end date
  - Type (outpatient, office, inpatient)
  - Care Site\*
  - Provider\*
- Provider
  - Gender\*
  - Specialty\*
- Care Site
  - Type (e.g. office, hospital, ER, nursing home)
- Payer Plan Period
  - Start/End Date
  - Type (e.g. Medicare, Medicaid)

*\*: Many missing or “other” entries*





# OMOP/N3C Data Tables



- Procedures
  - (e.g. Ultrasound, ECG, critical care, MRI)
  - Date
- Devices
  - (e.g. Stent, Catheter)
  - Start/End date
- Drugs
  - (e.g. iron, cefazolin 1000 mg, ...)
  - Start/End date
  - Route\* (oral, intravenous, ...)
- Measurements/Labs
  - (e.g. BMI, Heart Rate, COVID PCR test, ....)
  - Date
  - Value (e.g. 21.6, 72 bpm, “positive”)
- Conditions/Diagnoses
  - Name+ID (e.g. Pneumonia, Fever, Otagia, Hypothyroidism, ....)
  - Start/End date
- Death\*\*\*
  - Date\*
  - Type\* (EHR record, Gov Record, ...)
  - Cause\* (Sepsis, COVID-19, ...)

*\*: Many missing or “other” entries*



# External/Supplemental Data (Sample)



- Social Deprivation Index
  - By State and County
  - Population
  - % high needs population
  - % single parent
  - % unemployed
  - ...
- Healthdata.gov COVID Policies
  - State, County
  - Policy Date Start/End
  - Policy Type (“Phase 2”, “Shelter in Place”, “Food and Drink”, ...)
  - Policy details (text)
- SafeGraph Anonymized US GPS Movement Data
  - By census block
  - # Devices
  - Median & mean distance travelled
  - Destination census blocks
  - Time spent at home/in travel/etc
- Rural-Urban Commuting Area (RUCA) Codes
  - Zip code
  - Primary RUCA (10 categories, Metro core, small town, rural, ...)
  - Secondary RUCA (categorizing amount of commute to more urban areas)





## On Demand Learning Resources:

- Enclave Training Portal!
- Book of OHDSI: [book.ohdsi.org](https://book.ohdsi.org)
- EHDEN Academy: [academy.ehden.eu](https://academy.ehden.eu)
- OHDSI Tutorials: <https://www.ohdsi.org/past-events/>
- Data Tables Overview:  
[https://ohdsi.github.io/CommonDataModel/cdm531.html#Clinical\\_Data\\_Tables](https://ohdsi.github.io/CommonDataModel/cdm531.html#Clinical_Data_Tables)

## *Join the Conversation!*

Ask questions on the OHDSI Forums ([forums.ohdsi.org](https://forums.ohdsi.org)) or Twitter ([@OHDSI](https://twitter.com/OHDSI))



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

N3C  OMOP



NATIONAL CENTER  
FOR DATA TO HEALTH

## Harmonization is complicated!

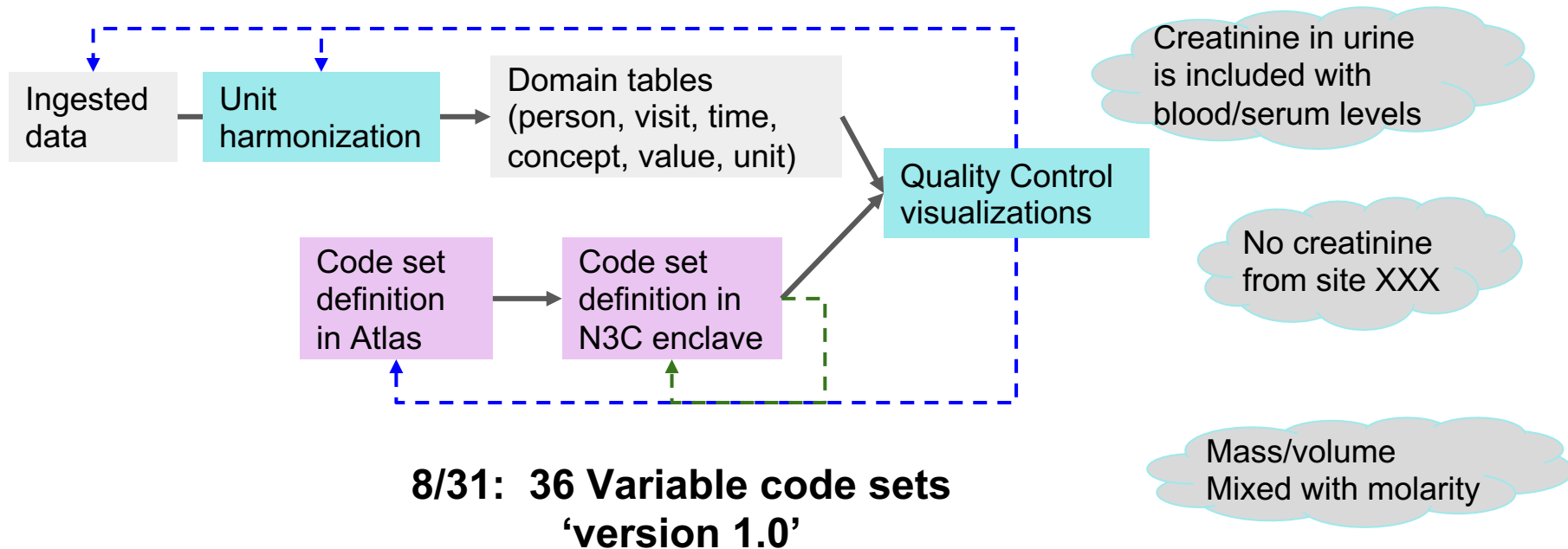
Creatinine in urine  
is included with  
blood/serum levels

No creatinine  
from site XXX

Mass/volume  
Mixed with molarity



## Harmonization is complicated!





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# Death Data?



NATIONAL CENTER  
FOR DATA TO HEALTH

Source data from partners don't include exactly the same information...



Death data supported?	Y
Death date required?	Y
Death cause supported?	Y
Discharge disposition supported?	Y



Death data supported?	Y
Death date required?	N
Death cause supported?	Y
Discharge disposition supported?	Y



Death data supported?	Y
Death date required?	Y
Death cause supported?	N
Discharge disposition supported?	N

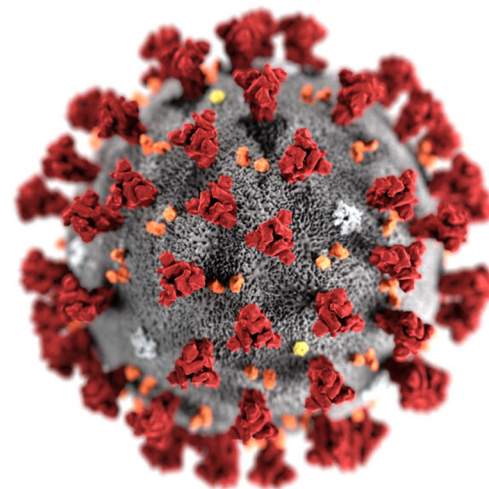


Death data supported?	Y
Death date required?	N
Death cause supported?	N
Discharge disposition supported?	N

Emily Pfaff, PhD, UNC



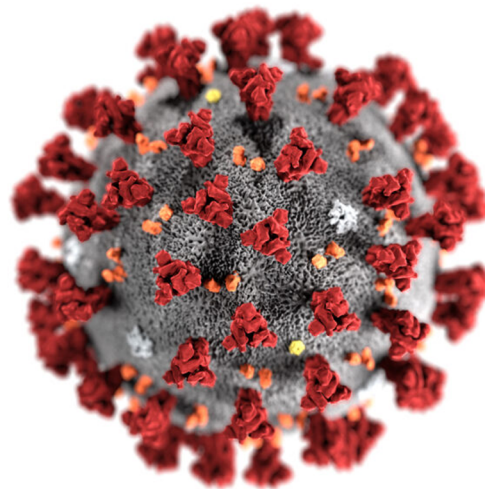
- **Multiple ways to identify COVID+ cases**
  - Many test types (some PCR, some Antibody)
  - Qualitative results encoding (“Elevated”, “High”, “Present”, ...)
  - Variable expressions of specificity (“Presumed positive / negative”)
  - ICD10: U07.1 (not used regularly in early 2020)







- **Multiple ways to identify COVID+ cases**
  - Many test types (some PCR, some Antibody)
  - Qualitative results encoding (“Elevated”, “High”, “Present”, ...)
  - Variable expressions of specificity (“Presumed positive / negative”)
  - ICD10: U07.1 (not used regularly in early 2020)
- **N3C-Supported Shared Definitions (via code templates & concept sets)**





# A Straw Poll



Element	Priority	Difficulty
Death status	2.93	2.22
Inpatient admission	2.92	2.73
ICU admission	2.86	2.36
Ventilator use	2.86	2
Supplemental oxygen	2.64	1.55
Vaccine administration	2.64	1.57
Vaccine adverse events	2.38	0.78
Ventilator settings	2.29	1.18
Insurance status	2.23	2.3
ICU flowsheet	2.14	1.11
Convalescent plasma	1.92	1.44
Pregnancy status	1.85	2
Family relationship	1.69	0.6
Other blood products	1.69	1.8
EKGs	1.62	1.44
DNR status	1.62	1.33

*January 11, 2021, "N3C Data Elements & CDMs"  
Symposium*

## Coding Key

N3C Clinical Domain Teams

**Priority** (n= 13)

3 = High

2 = Medium

1 = Low

Contributing CTSA Hubs

**Difficulty** (n=10)

3 = Easy

2 = Average

1 = Hard

-1 = Not Possible



National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

## Site-Specific Info



NATIONAL CENTER  
FOR DATA TO HEALTH

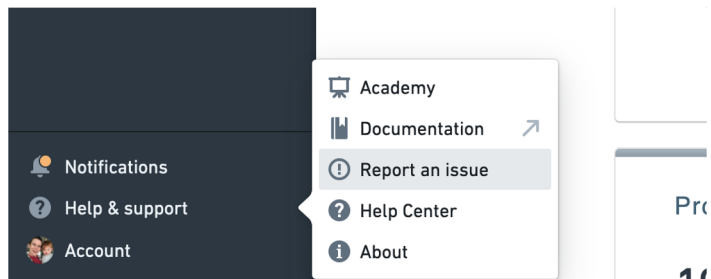
- **All tables have a *data\_partner\_id* column**
  - Indicates which data partner every row comes from
  - Numeric identifier: data partner 732 is **redacted** .



# Site-Specific Info



- **All tables have a *data\_partner\_id* column**
  - Indicates which data partner every row comes from
  - Numeric identifier: data partner 732 is **redacted**.
- **Known persistent issues**
  - <https://unite.nih.gov/workspace/documentation/product/n3c-info/data-issues>
- **Found something strange?**  
Report an Issue under  
“Help & Support” in the Enclave!





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C Community



NATIONAL CENTER  
FOR DATA TO HEALTH

## N3C is more than data, it's a community!

- Registered Users: **2200+**
- Enclave Users: **1300+**
- Research Projects: **160+**
  
- Institutions with Data Usage Agreements: **190+**
- Data Transfer Agreement Signatories: **85+**





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C Community



NATIONAL CENTER  
FOR DATA TO HEALTH

## N3C is more than data, it's a community!

- Registered Users: **2200+**
- Enclave Users: **1300+**
- Research Projects: **160+**
- Institutions with Data Usage Agreements: **190+**
- Data Transfer Agreement Signatories: **85+**
- Slack channels: **30+** <https://cd2h.slack.com>
- Google groups: **30+**





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C Community



NATIONAL CENTER  
FOR DATA TO HEALTH

## N3C is more than data, it's a community!

- Registered Users: **2200+**
- Enclave Users: **1300+**
- Research Projects: **160+**
- Institutions with Data Usage Agreements: **190+**
- Data Transfer Agreement Signatories: **85+**
- Slack channels: **30+** <https://cd2h.slack.com>
- Google groups: **30+**
- Regular meetings: **20+**  
<https://covid.cd2h.org/n3c-calendar>





National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C 'Map' (Unofficial, Incomplete)



NATIONAL CENTER  
FOR DATA TO HEALTH

Data Partnership &  
Governance

Synthetic Data

Collaborative Analytics

Phenotype & Data  
Acquisition

Data Ingestion &  
Harmonization





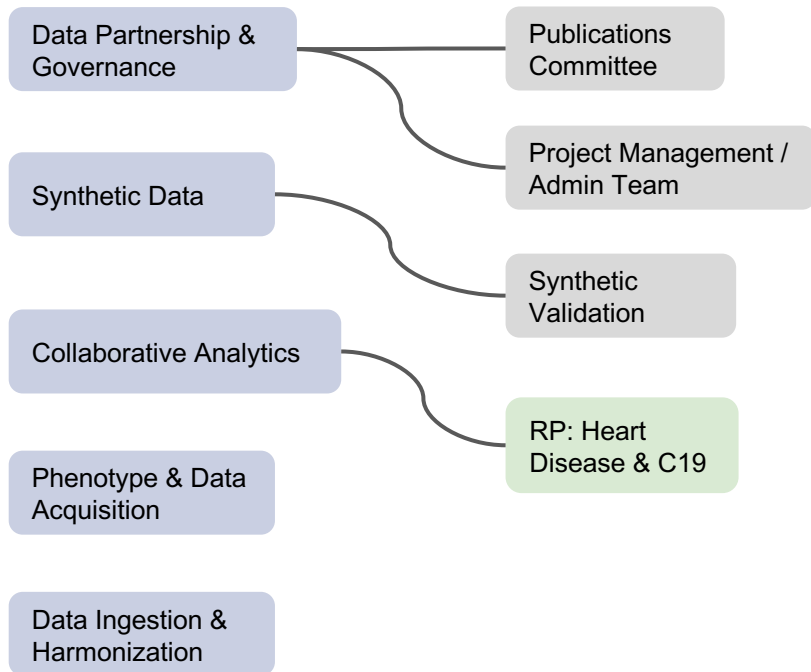
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C 'Map' (Unofficial, Incomplete)



NATIONAL CENTER  
FOR DATA TO HEALTH





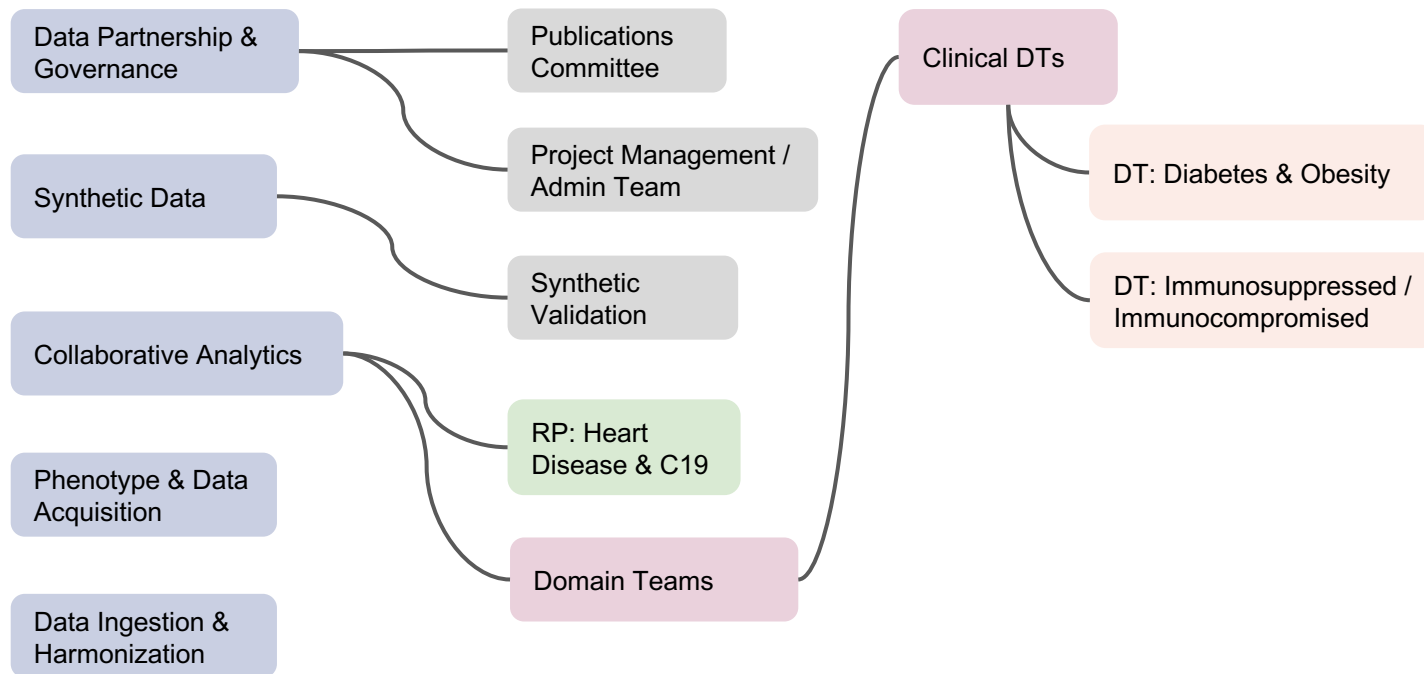
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C 'Map' (Unofficial, Incomplete)



NATIONAL CENTER  
FOR DATA TO HEALTH



<https://covid.cd2h.org/domain-teams>



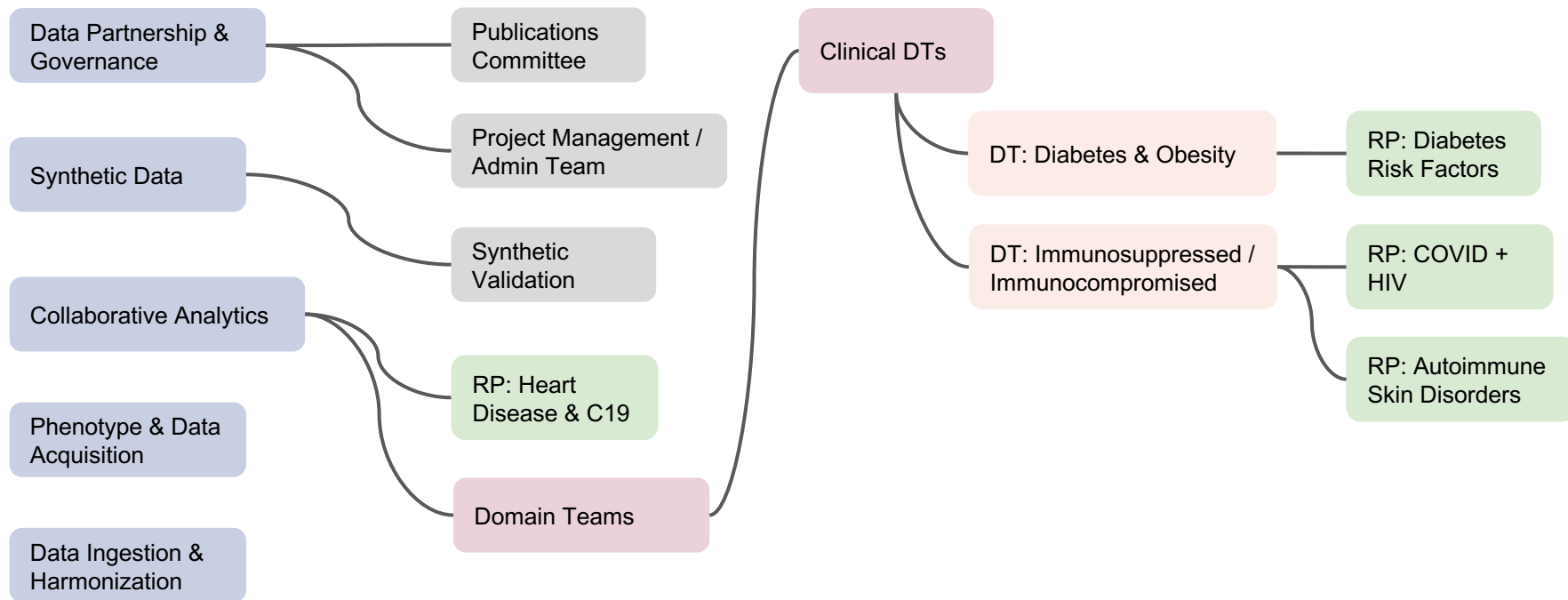
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C 'Map' (Unofficial, Incomplete)



NATIONAL CENTER  
FOR DATA TO HEALTH



<https://covid.cd2h.org/domain-teams>



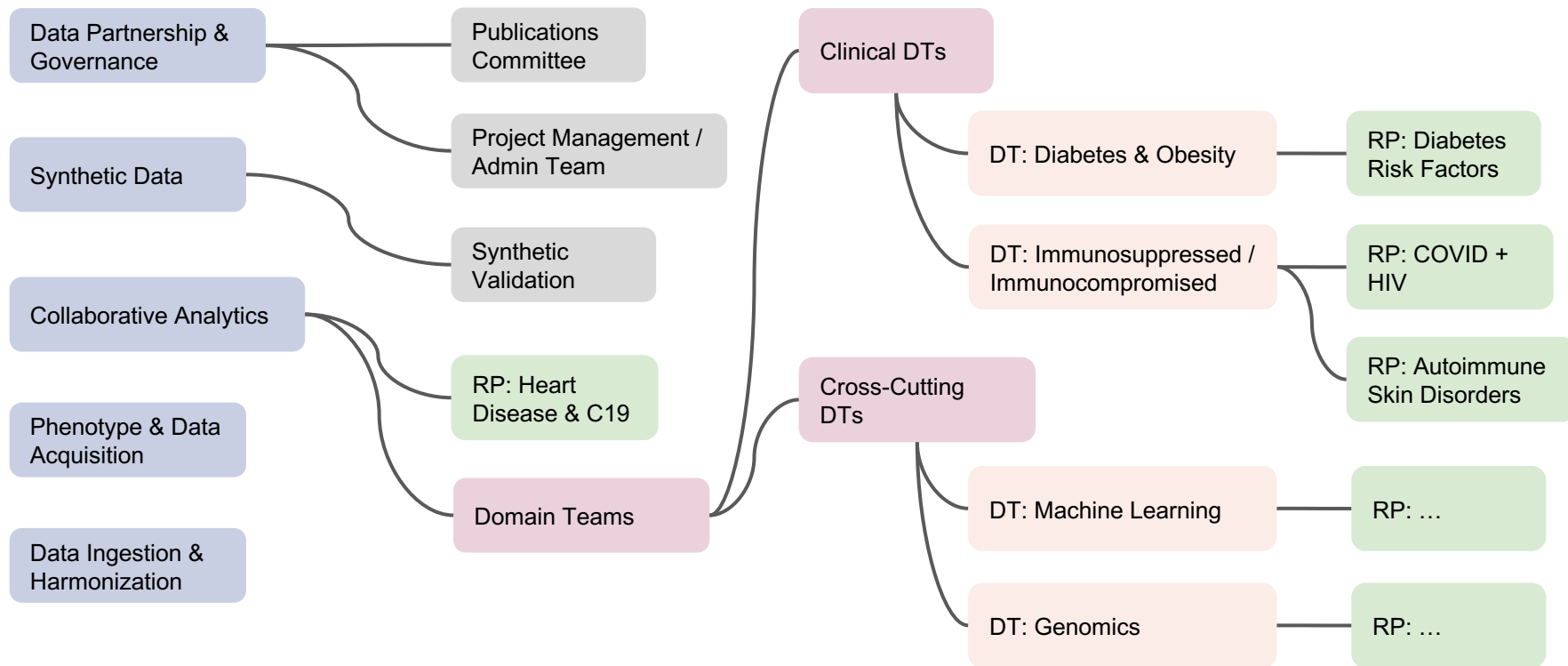
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C 'Map' (Unofficial, Incomplete)



NATIONAL CENTER  
FOR DATA TO HEALTH



<https://covid.cd2h.org/domain-teams>



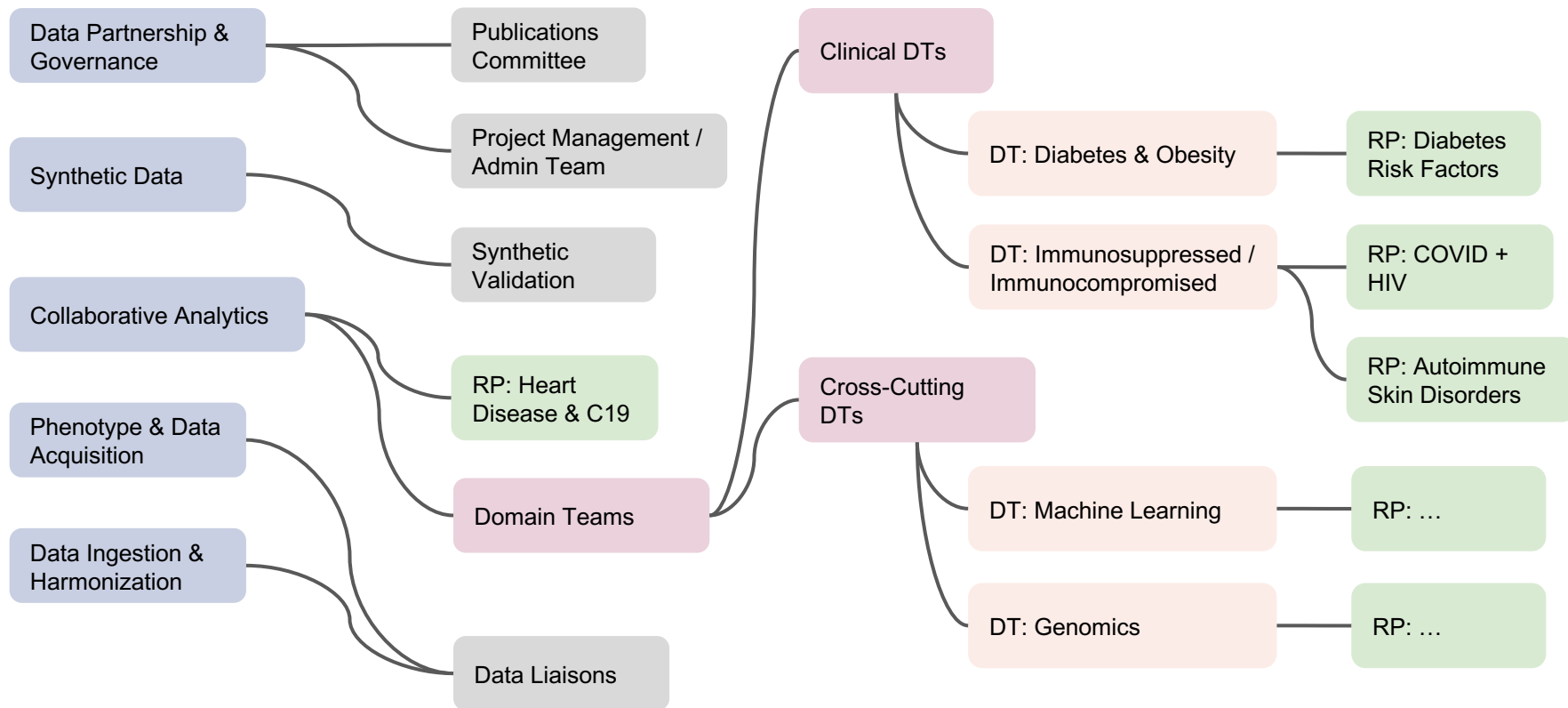
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C 'Map' (Unofficial, Incomplete)



NATIONAL CENTER  
FOR DATA TO HEALTH



<https://covid.cd2h.org/domain-teams>



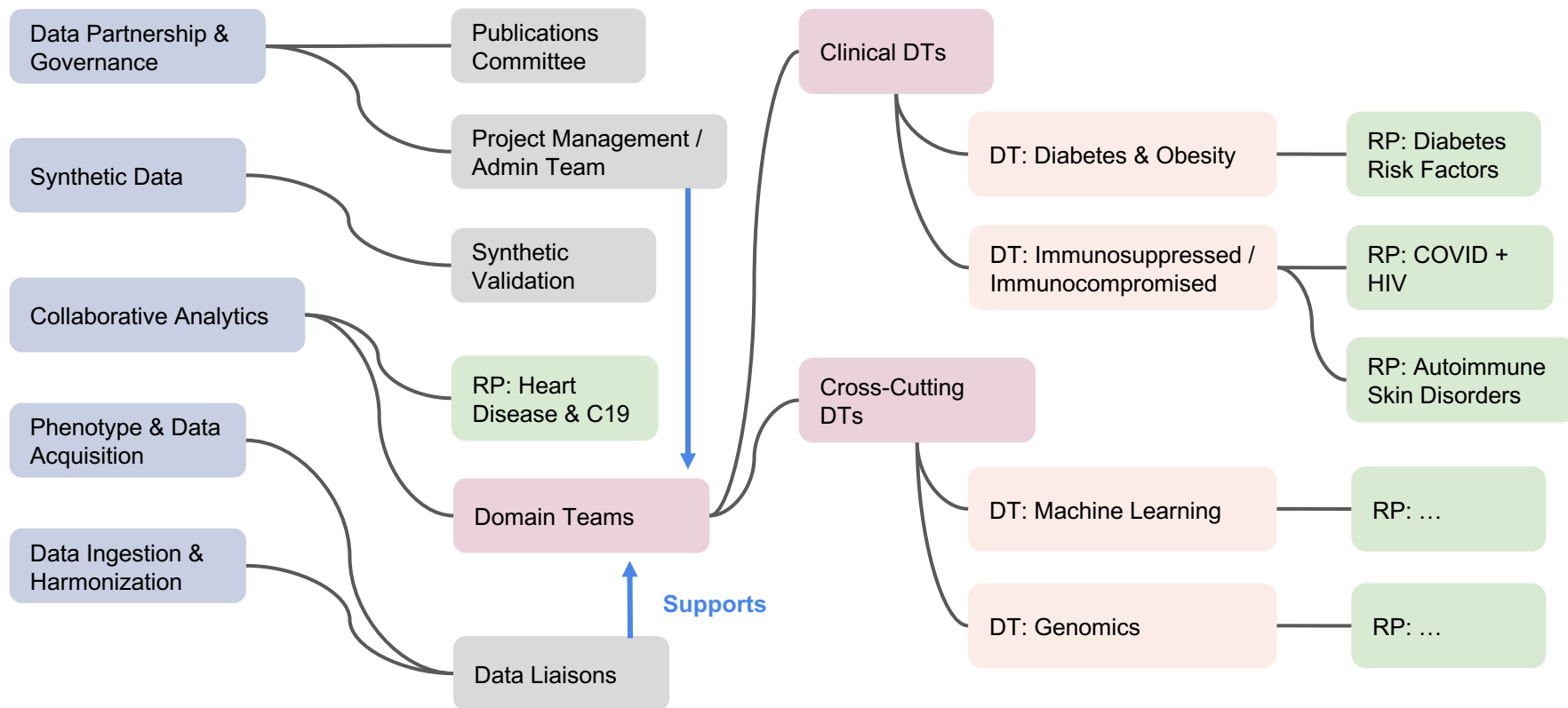
National  
COVID  
Cohort  
Collaborative

A program of NIH's National Center  
for Advancing Translational Sciences

# N3C 'Map' (Unofficial, Incomplete)



NATIONAL CENTER  
FOR DATA TO HEALTH



<https://covid.cd2h.org/domain-teams>



# Next Steps



- Check out the resources discussed today
- Register (if you haven't already)
  - <https://covid.cd2h.org/for-researchers>
- Visit the N3C website, especially Domain Teams
  - <https://covid.cd2h.org/domain-teams>
- Access the Enclave and browse around, especially training resources
- Most of what I've covered here today is in an Orientation A Session.
- There is also an Orientation Session B, where N3C staff introduce analysis techniques and tools for working with N3C data effectively!



National  
COVID  
Cohort  
Collaborative