National COVID Cohort Collaborative (N3C)

Sharon Patrick, National COVID Cohort Collaborative Coordinator

Lesley Cottrell, PhD – Director: West Virginia University UCEDD
Objectives

- Define and describe the N3C database;
- Review embedded N3C variables;
- Discuss N3C implications for AUCD; and
- Review next steps
What is the National COVID Cohort Collaborative
N3C Enclave: The largest limited EHR data set in US history

- Billions Rows of Data
- Millions of COVID Cases
- Millions of People
- Hospitals and Clinics Across the Country
- Procedures, Visits, Medications: Goes back 2 years before COVID Test
N3C Literature Examples

- Database rationale, design, structure
- Mechanisms associated with COVID-19 severity
- Early severity prediction
- Health disparities and equity of COVID-19
- Deep phenotyping of hospitalized COVID-19 patients
- Treatment regimen patterns and trends
Translational Research Continuum

T1TR

Basic Research
Pre-Clinical Studies
Clinical Efficacy
Clinical Effectiveness

T2TR

Dissemination
Implementation
Impact
N3C Data and Structure
External Datasets

Users can view the list of external datasets that have been imported into the Enclave, as well as the datasets requested and their current status in the review process by viewing the External Dataset Registry

The National COVID Cohort Collaborative (N3C)
A National Partnership to Fight COVID-19

The N3C aims to improve the efficiency and accessibility of analyses using a very large patient-level COVID-19 clinical dataset and demonstrate a novel approach for collaborative pandemic data sharing.

VALUE TO ORGANIZATIONS:
- Access to large scale COVID-19 data from across the nation
- Sample data for grant proposals
- Team science opportunities for new questions and to test informatics methods
- Access to Domain Teams, statistics, machine learning (ML), informatics expertise
- Training on ML analytics, NLP methods, tools, software, additional datasets

WAYS TO ENGAGE:
- **Join a Domain Team:** Acute Kidney Injury (AKI), Applicable Data Methods Standards (ADMS), Critical Care/Ventilation/Multi-organ Dysfunction Syndrome (MODS), Diabetes & Obesity, Elder Impact, Emergency Services, Imaging, Immunosuppressed/compromised, Oncology, Pediatrics, Pharmacoepidemiology, Pharma-commercial, Pregnancy, Short/long-term Complications, Social Determinants of Health (SDoH)... or launch a new one
- Apply for access to the N3C Data Enclave

Join today at Covid.cd2h.org
West Virginia Clinical and Translational Science Institute

Domain Teams & Common Resources

**Domain Teams**
N3C Domain Teams enable researchers with shared clinical questions surrounding COVID-19 to collectively analyze data within the N3C Data Enclave and collaborate more efficiently in a team science environment.

Learn More

**FAQs**
Explore frequently asked questions specifically related to the access and flow of data in the N3C Data Enclave.

Learn More

**Projects**
See a list of N3C projects that have been submitted through the Data Use Request (DUR) process and approved by the Data Access Committee (DAC).

Learn More

**Support**
If you have any questions, problems, or concerns, please visit the N3C Support Desk to see available support options.

Learn More

**Training**
Watch tutorials to learn about the N3C Data Enclave.

Learn More

**Tools**
Learn about the tools included within the N3C Data Enclave. A searchable tool registry will be available soon!

Learn More
N3C Domain Teams enable researchers with shared interests analyze data within the N3C Data Enclave and collaborate more efficiently in a team science environment. These teams provide an opportunity to collect pilot data for grant submissions, train algorithms on larger datasets, inform clinical trial design, learn how to use tools for large scale COVID-19 data, and validate results.

Domain Teams are enabled by Slack channels for discussion, meetings, and document management and are supported by N3C workstreams. N3C encourages researchers of all levels to join a Domain Team that represents their interests, or to suggest new clinical areas to explore. A Domain Team can submit one or more research projects, but collaboration is encouraged for similar concepts.
Lifespan Disability Domain

Mission: To advance practices and policies that improve the COVID-19 education, access, and health outcomes of all people with developmental and other disabilities as well as their families, by examining extent to which individuals with disabilities have been affected by COVID-19 diagnoses and access to appropriate treatments and follow up.

Purpose: To establish a group of experts who will examine the N3C data in terms of COVID's impact on individuals (of all ages) who have at least one disability.

- Disability types would include: intellectual, developmental, vision impairment, deaf or hard of hearing, mental health conditions, acquired brain injury, and mobility/physical disabilities.

https://covid.cd2h.org/disabilities
Initial Research Question

• What disparities, if any, exist based on disability related to:
  • Testing
  • COVID diagnosis
  • Treatment options
  • Prognosis
Domain Recording
West Virginia Clinical and Translational Science Institute

N3C Registration/Training

https://covid.cd2h.org/Tutorials

Registration for Documents, Meetings & the N3C Data Enclave

Requires Authentication

Enclave Checklist

Training Office Hours:
Tuesdays & Thursdays at 10-11 am PT/1-2 pm ET
Registration Required at this link

Additional Training Tutorials available in the Enclave
N3C Enclave Orientation Sessions

https://covid.cd2h.org/training

Session A - Recommended before attending Session B
This session is for those who want to learn about N3C, as well as how to engage with project teams and access the data.

Learning objectives:
- Provide a general overview of N3C, including goals, organization, and community resources such as Domain Teams and Data Liaisons.
- Introduce the 3 data tiers available and important considerations for research driven by the data harmonization process.
- Discuss resources for training and support and the process of the Data Use Request (DUR) required for researcher access.

Session B - Completion of Session A is recommended
This session is for analysts, statisticians, data scientists, or anyone who wants to gain a broader understanding of the tools needed to work with the data.

Learning objectives:
- Focus on technical aspects of working with data in the secure N3C Enclave, including use of OMOP concept sets and N3C-specific tooling, such as the Concept Set Browser.
- Introduce commonly used analysis tools, such as Contour and Code Workbooks and corresponding workflows for simple analyses.
- Introduce the Enclave Knowledge Store, a mechanism for sharing and using community developed code and data across projects.

Alternating Tuesdays | 8-9:30am PT/11am-12:30pm ET
Register for Part A

Alternating Tuesdays | 8-9:30am PT/11am-12:30pm ET
Register for Part B

For users of all types to learn how to use the N3C Data Enclave.

Sessions will be ongoing and will alternate week by week.
N3C Participation: Requirements, Process, and Access
First Time User Prerequisites

- Signed DUA
- HSP / Security Training
- Register at N3C
A Data Use Agreement is required to outline the terms of use for the Enclave portal as well as govern policies around the data contained within. A Data Use Agreement must be in place for your institution, or you (Citizen Scientist) before you will be able to login to the N3C portal. A Data Use Agreement is executed by NCATS and a research institution (or directly with a researcher in the case of a Citizen Scientist who is not affiliated with an institution).

**DUA Process**
- Check to see if institution is listed at [https://covid.cd2h.org/duas](https://covid.cd2h.org/duas)
Goal of the Data Use Agreement is Privacy Protection to Promote broad access:

- COVID-Related research only
- No re-identification of individuals or data source
- No download or capture of raw data
- Open platform to all researchers
- Security: Activities in the N3C Data Enclave are recorded and can be audited
- Disclosure of research results to the N3C Data Enclave for the public good
N3C Account Creation

Purpose: Your N3C account grants you access to 2 distinct areas:

- N3C Data Enclave
- N3C Team Drive [which includes Google Groups, and Slack]

You may begin the registration process at: https://labs.cd2h.org/registration
How to Register

Start here: https://covid.cd2h.org/
How to Register

You will be redirected here: https://covid.cd2h.org/N3C_data_enclave

First time user? The Enclave Registration Checklist to streamline your registration

Already registered? Log in here
N3C Account Creation

N3C Data Enclave account requirements:

• Confirmation that an executed Data Use Agreement (DUA) exists between your organization and NCATS.
• Completion of NIH Information Security and Information Management Training
  • Each individual data user must complete the following course: 2020 Information Security consisting of these 5 modules: Information Security, Counterintelligence & Insider Threat, Privacy, Records Management, Emergency Preparedness. (Allow 60-90 minutes to complete all 5 modules.)
  • After completing the 5th module of the 2020 Information Security course, a button will appear on the top bar that says "Print Certificate".
• Documentation of completion of Human Subjects Research Protection training in alignment with your institution's policies if you intend to access the de-identified or limited datasets. (This training requirement does not apply to investigators who will only access the synthetic dataset.)
  • e.g. CITI Training (Human Subjects Protection Training Courses)
    • Good Clinical Practice (GCP)
    • Health Information Privacy and Security (HIPS)
    • Responsible Conduct of Research Courses (RCR)

N3C registration is open to all individuals. Please note there are 2 different registration paths depending on institutional affiliation

Details: [https://covid.cd2h.org/account-instructions](https://covid.cd2h.org/account-instructions)
InCommon Account:

- An InCommon account is available to any user who belongs to an institution that is a member of the InCommon Federation, a federated login system used by N3C.
- Check here: https://www.incommon.org/community-organizations to see if your institution participates in InCommon. If your institution is not listed, you will need to sign up with a Login.gov account (see next page).
- During registration select "Yes" to the enclave access suggestion.
- If you wish to access N3C slack or Google groups select the communications methods you wish to use and enter your slack and/or Google-enabled email addresses. (They don't have to be the same as your institutional email.)
- Access to the Enclave will be enabled after your account has been verified to be covered by a DUA that matches your account.
- **Please allow 5 business days for account creation.**
- **Note:** If your institution does not participate in InCommon, you can get more information about joining the InCommon Federation here: https://www.incommon.org/federation/federation-join/
N3C Account Creation
cont.

**Login.gov Account:**

- A Login.gov account is used by many different US Government related services and will also allow you access to the N3C Data Enclave when (1) either your institution does not participate in InCommon or (2) you are a Citizen Scientist not working on behalf of an institution.

- Choose a "Registration Email" - this should be if at all possible an email associated with some organization like your workplace or school, but NOT an email associated with any institution listed at [https://www.incommon.org/community-organizations/](https://www.incommon.org/community-organizations/)

- Create (or use an existing) Account at [https://login.gov](https://login.gov) using your Registration Email.
  - **Citizen Scientists** (not affiliated with an institution) must use the same email address to register for an account as was used to establish the DUA.

- Register with N3C at [https://labs.cd2h.org/registration/](https://labs.cd2h.org/registration/)
  - When completing the registration, use your Login.gov account and select "Yes" to the Enclave access suggestion.
  - If you wish to access the N3C Slack platform or Google groups, select the communications methods you wish to use and enter your Slack and/or Google-enabled email addresses. (They don't have to be the same as your registration email used with Login.gov).
  - Access to the Enclave will be enabled after your account is verified to be covered by a DUA matching your account.
Understanding the Data Use Request (DUR)
A Data Use Request (DUR) is an online application submitted by the researcher within the N3C Data Enclave that describes the scope and nature of the project and justifies the requested data access tier. Approval of the DUR by the Data Access Committee (DAC) is required before a researcher can access data within the enclave. (Approved DURs are valid for 1 year.)

- If you intend to request access to the limited dataset (LDS), you will need to have obtained an IRB determination letter from your institution.
- The Lead Investigator should submit the Data Use Request online within the N3C Data Enclave located at: https://unite.nih.gov/workspace/slate/documents/data-use-request/latest
- HELPFUL Hint- Lead investigator reviews and approves all Collaborators who want to join the study. Once the Collaborator has been approved it will be sent to the DAC for review and approval
- Others can request to join an existing DUR-search projects within the enclave
- **Please allow 21 days for DAC approval**
- After the DUR has been approved a workspace within the enclave will be provisioned.
# Data Use Request Form: Project & Collaborator Information

## Project Information
- Enter project title: public
- Enter project abstract: public
- Enter project rationale: private
- Indicate decision to allow contributors

## Collaborator Information
- Collaborator name
- Collaborator institution
- Collaborator email address

### 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.

<table>
<thead>
<tr>
<th>Collaborator's First Name</th>
<th>Collaborator's Last Name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- [ ] This collaborator belongs to my institution

<table>
<thead>
<tr>
<th>Collaborator Email</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

- Add Additional Collaborator

---

Accessing Institution:

Oregon Health & Science University

Project Title (Public):

Please enter your new project title.

Research Project Abstract (Public):

Research Project Rationale (Confidential to N3C DAC):

Allow other researchers to join this project?

- Allow
- Do not allow
Approved Enclave Projects

https://covid.cd2h.org/projects

View the list of N3C Data Enclave Projects that have been approved by the (DAC).

**Projects:**

**Acute Kidney Injury in Pediatric COVID-19 Patients**

Adult patients with COVID-19 have a high incidence of acute kidney injury (AKI). The cause of AKI is multifactorial and is hypothesized to include direct viral injury and inflammation of the kidneys, pre-existing conditions conferring a higher risk, and treatment sequelae such as nephrotoxic medications. The purpose of this study is to better characterize acute kidney injury in pediatric patients (> 18 years old) with and without COVID-19. Specifically we request Level 2 (de-identified) data to ask the following questions: 1) What is the incidence of AKI among pediatric patients, both with and without COVID-19; 2) What risk factors predict the development of AKI in pediatric COVID-19 patients; 3) What are the outcomes of patients with and without AKI; 4) What is the time-course of AKI resolution; and 5) How are treatments, including medication and technology (e.g. CRRT or ECMO) associated with AKI progression or resolution. These final three questions will stratify AKI by KDIGO severity criteria.

**Lead Investigator:** ADAM DZIORNY
**Accessing Institution:** University of Rochester

**Analysis of Time Windows as Determining Factors for COVID-19 Patient Outcomes**

Protocol that determines if a patient should be tested for COVID-19 may depend on presence of certain symptoms and vital signs. If presence of certain symptoms in combination with the length of time before getting tested has an impact on patient outcomes, we would like to find the optimal time window that leads to better patient outcomes.

**Lead Investigator:** Arpi Patil
**Accessing Institution:** University of Virginia
West Virginia Clinical and Translational Science Institute

View Existing Projects & Request to Join

### Projects to Join

<table>
<thead>
<tr>
<th>Title</th>
<th>Nonconfidential Research Statement</th>
<th>Lead Investigator Name</th>
<th>Email Lead Investigator</th>
<th>Lead Investigator Accessing Institution</th>
<th>Request to Join Project as Collaborator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Test Aug 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Request to Join</td>
</tr>
<tr>
<td>Test Aug 16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Request to Join</td>
</tr>
<tr>
<td>Bugs bunny returns</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Request to Join</td>
</tr>
<tr>
<td>way</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Request to Join</td>
</tr>
<tr>
<td>Testing the System</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Request to Join</td>
</tr>
<tr>
<td>test test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Request to Join</td>
</tr>
<tr>
<td>DUR Test project</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Request to Join</td>
</tr>
<tr>
<td>Finding Name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Request to Join</td>
</tr>
</tbody>
</table>

### Projects to Explore

<table>
<thead>
<tr>
<th>Title</th>
<th>Nonconfidential Research Statement</th>
<th>Lead Investigator Name</th>
<th>Email Lead Investigator</th>
<th>Lead Investigator Accessing Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Test Project</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Data Ingestion and Harmonization

HSC operations to extract, transform, and load data from contributing sites into a common format. Additionally, the team will contribute to the cross-workstream efforts of normalize data elements into a common form, including units of measure and outlier tagging.
Need Assistance or Have a Question?

If you need assistance with the N3C Data Enclave, the following options are available:

- Attend Support Desk office hours on Tuesdays & Thursdays at 10-11 am PT/1-2 pm ET [Register here.]
- Submit a [Support Request](#).
- Consult the [Frequently Asked Questions (FAQs)](#).
- A quick start tour is available inside the Enclave.
- [Video tutorials will be uploaded as they are created.](#)
Attribution

N3C Administration and Leadership

Andrea Volz, N3C Communications Manager
Ken Gersing, NCATS-Co Leader N3C
Davera Gabriel, N3C Sr Research Terminologist
Julie McMurry, N3C Operational Data Manager
Anita Walden, N3C Assistant Director
Thank you!

PARTNERING FOR BETTER HEALTH

West Virginia Clinical and Translational Science Institute

@WVCTSI  Find us on  www.wvctsi.org