

# The role of government in data-sharing, an NIH Perspective

Penny Wung Burgoon, PhD

*Office of Policy, Communications and Education  
National Center for Advancing Translational Sciences, NIH*

# Data Sharing and the government perspective

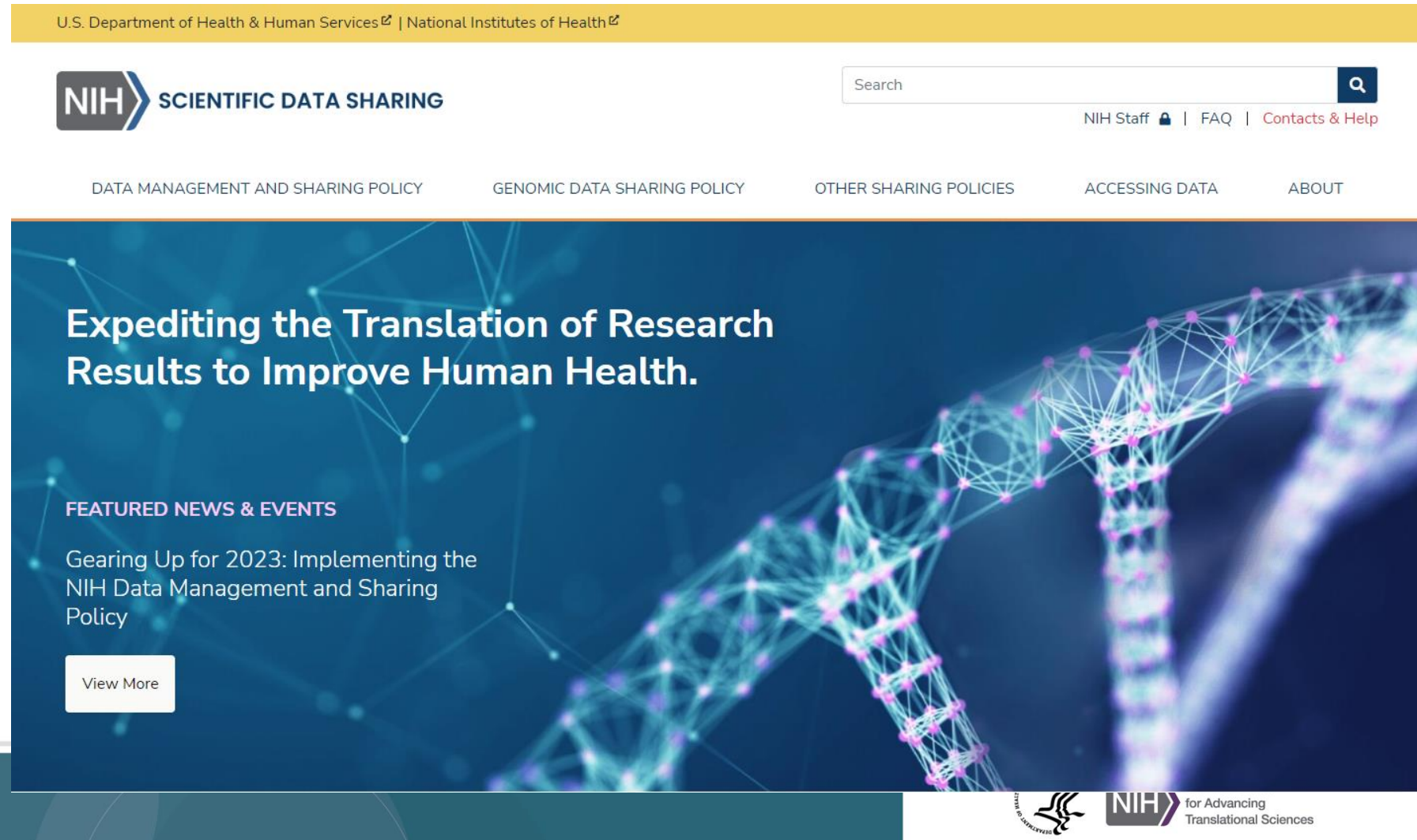
- NIH Policy for Data Management and Sharing – where to go for updates and resources
- The NCATS National COVID Cohort Collaborative, N3C



# Questions about the NIH Data sharing policy?

## [Home Page | Data Sharing \(nih.gov\)](#)

- Provides a central source of guidance related to multiple NIH data sharing policies
- Covers Data Management and Sharing, Genomic Data Sharing, Model Organisms, and Research Tools policies
- Content will be updated





## NIH Policy for Data Management and Sharing

- **Submission of Data Management & Sharing Plan for all NIH-funded research** (*how/where/when*)
- **Compliance with the ICO-approved Plan** (*may affect future funding*)
- **Effective January 25, 2023** (*replaced 2003 Data Sharing Policy*)



# NCATS National COVID Cohort Collaborative (N3C): Electronic Health Record Data (EHR) to Study COVID-19

## N3C EHR Database

Diverse Set of De-Identified &  
Limited Data Set Clinical Health  
Information From Electronic Health  
Records



Sites: 82



Persons: 20.7 million



COVID+ Cases: 8,153,703



# of Rows: 27.7 billion



Clinical Observations: 2.6 billion



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

## Overarching Purpose and Duration

- To understand clinical course of COVID-19
- A 5-year program with the ability to extend with renewed agreements
- New submissions collect data from Jan. 2018-present, and refreshed every 2-4 weeks

## N3C Database Controlled Access

- Institutions Sign Data Use Agreement
- NIH Data Access Committee (DAC) Reviews data use requests
- Data cannot be removed
- Adheres to Federal and NIH policies & regulations
- Extensive learning, training, and community resources available

## Answering Important Public Health Questions

>70 preprints or publications  
>1000 Google citations

What are  
COVID-19  
symptoms?

What are risk  
factors for  
severe COVID?

What are  
risk factors  
for Long  
COVID?

What are best  
treatments for  
COVID-19?

Partners &  
Collaborators:

NCATS \* NIGMS \* CTSAs \* IDeA-CTRs \* NLM \* ODSS \* NIAID \* NIBIB \* NICHD \* NHLBI \* THRO \* FDA \* BARDA \* CMS \* ONC, CD2H,  
Palantir \* Datavant \* Acumen \* SAMVIT \* OCHIN \* Regenstreif \* MDClone \* Microsoft \* Sage Bionetworks



NIH National Center  
for Advancing  
Translational Sciences



# NCATS National COVID Cohort Collaborative (N3C): Rapid Access to Real-World Data Informs Urgent COVID-19 Questions

## PASC/Long-COVID Risk

- NCATS/RECOVER funded
- Used AI/ML to identify PASC before U09.9 code
- **Difficulty breathing associated with higher risk of PASC**
- **Vaccination associated with lower risk of PASC**



## RWE for mAb Effectiveness Across Variants

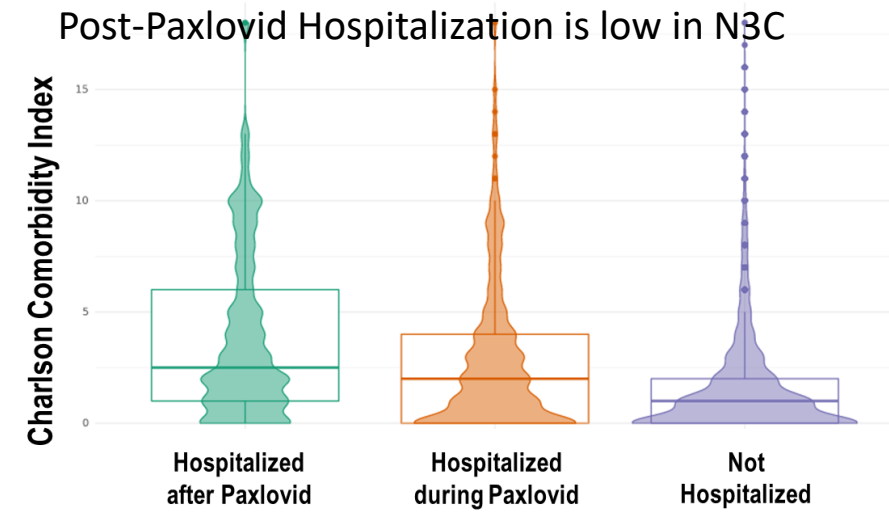
- Effectiveness of mAb treatments
- Viral Variants and impact on interventions
- **Multiple mAbs, mostly Delta variant**
- **89% reduction in 28-day mortality**



## RWE Analyses on Paxlovid Rebound: Unpublished, prepared for White House COVID Response Team

### RWE "Clinical Trials" in N3C

- At-a-glance counts to help power studies
- Define cases/controls
- Define machine readable EHR outcomes
- Limitations: Cannot replace randomized clinical trial studies, outcomes are less refined, missingness



Many similar questions being asked of NCATS/N3C data; coordination is key



NIH National Center  
for Advancing  
Translational Sciences