

# Data Science and Teamwork for Understanding PASC: PCORnet™ RECOVER EHR

Mollie R. Cummins, PhD, RN, FAAN, FACMI; Ramkiran Gouripeddi, MBBS, MS, FAMIA, Jackson Barlocker, Sukrut Shishupal, MS, Mary Beth Scholand, MD

## Background

- RECOVER is a large-scale research initiative from the National Institutes of Health (NIH), to understand, prevent, and treat Post-Acute Sequelae of SARS-CoV-2 infection (PASC), including “long COVID”.<sup>1</sup>
- University of Utah is a participating site in **RECOVER EHR**, a nationwide study focused on leveraging a national sample of electronic health record data for insight into PASC (figure 1).

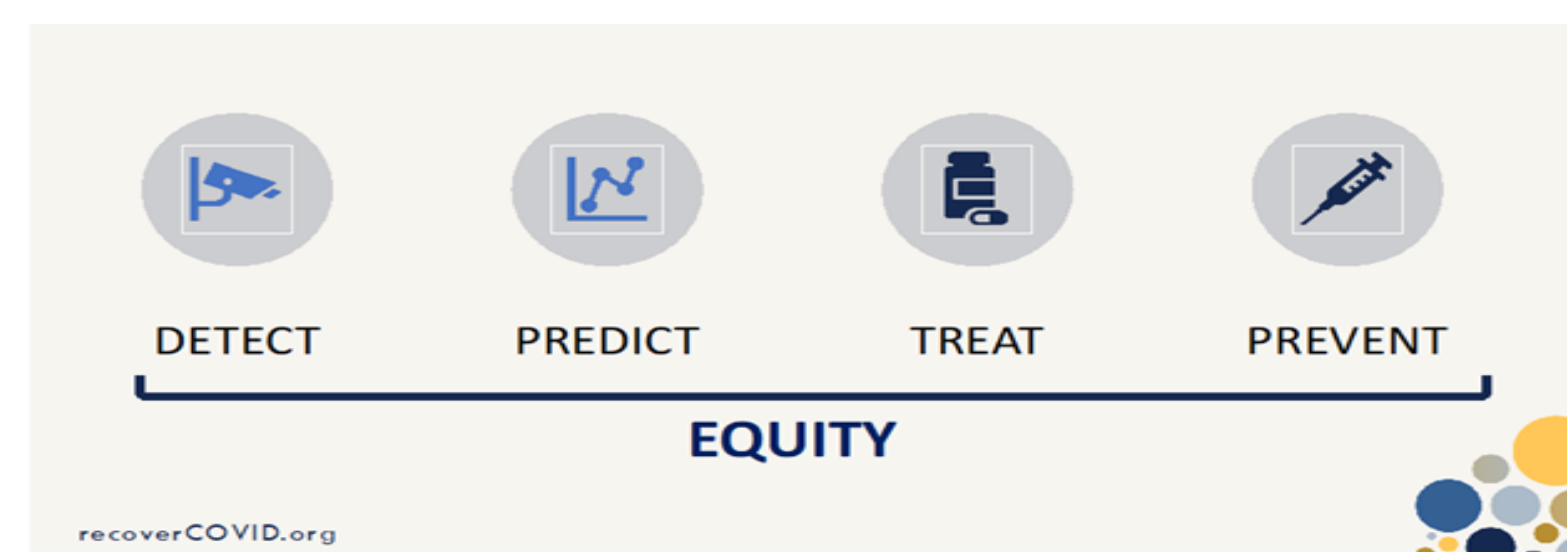


Figure 1. Scope of NIH RECOVER

## Methods

- Develop a single, unified repository of research data related to PASC that includes data from 40 sites nationwide – adults and children (figure 2), leverages existing PCORnet™ clinical research network infrastructure and CDM (common data model) for rapid outcomes.
- Use methods of data science, including artificial intelligence and machine learning, to create screening definitions (computable phenotypes and sub-phenotypes) and answer fundamental questions:
  - Subtypes of/ symptom clusters in PASC
  - Risk, mitigating, and protective factors for developing PASC
  - Predictive models for developing PASC
  - Assess racial, ethnic and socioeconomic disparities in people at risk for and living with PASC
  - Health care utilization of people living with PASC

## Reference

- RECOVER. RECOVER: Researching COVID to enhance recovery. National Institutes of Health. Accessed April 27, 2022, <https://recovercovid.org/>



## Results

Year 1 efforts focused on phenotype development and validation, disparities and risk factors (see table 1). Year 2 efforts will focus on additional data queries, support for planning and implementing clinical trials, target trial emulation, data linkage, and dissemination.

Table 1. Analyses during Year One

Topic	Focus	Q1	Q2	Q3	Q4
Detect	Phenotype development, refinement, validation				
	Characterize PASC				
	Iteratively quantify incidence and prevalence				
	Explore temporal trends				
Predict	Geographic, demographic, socioeconomic disparities				
	Examine risk factors				
Treat	Characterize treatments and patterns of therapeutic use				
	Therapeutic effectiveness				
Prevent	Vaccination linkage and quality improvement				
	Vaccine effectiveness				

- Large-scale cooperative effort, with early results available and more publications expected in coming months.
- Enabling production of computable phenotypes for use in RECOVER cohort studies focused on PASC
- Creates the capability to answer fundamental questions about the extent and nature of PASC
- Applications of large-scale EHR analytics for science and for policy-making/ health care planning

## Funding Acknowledgement

This research was, in part, funded by the National Institutes of Health (NIH) Agreement OT2HL161847-01. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the NIH.