National Center for Advancing Translational Sciences

National Institutes of Health (NIH) Tribal Consultation on the National COVID Cohort Collaborative (N3C) February 11, 2022, at 1:30 p.m. EST

This document is an overview for Tribal leaders and community members who are considering taking part in the upcoming **NIH Tribal Consultation on the NIH's National Center for Advancing Translational Sciences (NCATS) National COVID Cohort Collaborative (N3C)** web-based meeting on **February 11, 2022, at 1:30 p.m. EST.** A recorded informational webinar about the N3C Program is available for your review prior to the Tribal Consultation on the NCATS Tribal Consultation Event page, https://ncats.nih.gov/n3c/about/tribal-consultation.

National Institutes of Health (NIH)

NIH's mission is to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability.

- Each year, NIH invests about $41.7 billion[1] in medical research for the American people.
- NIH awards more than 80% of its funding through almost 50,000 competitive grants to more than 300,000 researchers at more than 2,500 universities, medical schools, and research institutions in every state.
- About 10% of NIH's budget supports projects conducted by nearly 6,000 scientists in its own laboratories, most of which are on the NIH campus in Bethesda, Maryland.
- NIH is part of the U.S. Department of Health and Human Services and is composed of the Office of the Director and 27 Institutes and Centers, many of which focus on specific organs or diseases. The Office of the Director is the central office at NIH that sets policy for NIH and for planning, managing, and coordinating the programs and activities of all the NIH components.

Purpose of Consultation

As NIH considers its research efforts in pursuit of its mission, it is critical that the agency conduct outreach and solicit input from Tribal Nations. NIH supports research projects that can increase our overall understanding of COVID-19 and its effects on American Indian or Alaska Native

> **American Indian and Alaska Native Data in N3C**
>
> Contributing sites provide race and ethnicity data to the National COVID Cohort Collaborative as part of a data set collected from patient electronic health records. Race and ethnicity information from health care settings is self-identified or assigned by health care providers and does not relate to Tribal enrollment and does not include Tribal affiliation.

(AI/AN) communities. NCATS is seeking input through Tribal Consultation to assist the NIH in

---

[1] This amount reflects the sum of discretionary budget authority of $40.3 billion received by NIH in FY 2020 under the "Further Consolidated Appropriations Act, 2020," Public Law (P.L.) 116-94, which includes $80 million for Superfund Research activities. The total also includes $1.23 billion derived from Public Health Services Evaluation financing and mandatory budget authority of $150 million for special type 1 diabetes research. The allocation of $225 million to NIH per P.L. 116-94 from the Nonrecurring Expenses Fund is excluded. Details regarding appropriations or budget requests are available from the Office of Budget.

deciding whether and how to make AI/AN data available to researchers through N3C for COVID-19 research.

COVID-19 Research Efforts

Since the beginning of the pandemic, health care providers and researchers have worked diligently to understand the novel SARS-CoV-2 virus and the disease it causes, COVID-19. Our understanding of COVID-19 — its signs, symptoms, effective courses of treatment — at the start of the pandemic was very limited. While this disease spread rapidly through communities, cities, and countries and crossed borders and oceans, important patient health information was being collected in many clinics and hospitals. Data collected in this manner are difficult to put together for researchers and doctors to understand the characteristics of this new disease. The medical and research communities urgently needed large amounts of data to better understand COVID-19, including how it spread, who was most at risk, which treatments helped, and what the effects of the disease were, including long-term effects.

In response to this urgent need, NCATS and its partners developed the N3C to collect existing electronic health record (EHR) data from hospitals and clinics and to make these data accessible to researchers seeking to understand COVID-19. The N3C effort is centered on the following:

1. Establishing a secure data repository for studying COVID-19-related data.
2. Receiving existing patient data derived from EHRs provided by participating U.S. health care sites.
3. Providing support for researchers using the N3C Data Enclave[2] to collaborate on COVID-19 research.
4. Ensuring research using the N3C Data Enclave follows the rules and expectations of NCATS and its partners to keep the data secure and protect patient privacy.

As of December 2021, the N3C has made data accessible to more than 3,000 researchers and clinicians to study the progression of COVID-19, identify risk and protective factors, search for effective treatments, understand the long-term disease effects, and determine how best to care for those with the disease. N3C is a collaborative partnership among many institutions and organizations to provide their clinical data in close to real time to improve our knowledge of COVID-19 and observe potential treatment strategies, while also keeping these patient data secure and used only as intended. N3C partners include *health care providers* (to provide the data in N3C), *NCATS* (to provide governance, oversight, and the secure Enclave to maintain and protect the data), and the *scientific community* and *research leaders with data science and clinical expertise* (to harmonize data so that it can be studied together and compared across the nation). Key partners of the N3C include the institutions of the NCATS Clinical and Translational Science Awards (CTSA) Program, the institutions supported by the NIH National Institute of General Medical Sciences' Institutional Development Award Networks for Clinical

---

[2] The N3C Data Enclave is a secure platform through which the harmonized clinical data provided by our contributing sites is stored. The data itself can be accessed only through a secure cloud portal hosted by NCATS and cannot be downloaded or removed.

and Translational Research (IDeA-CTR), and the NCATS CTSA Program's National Center for Data to Health (CD2H). N3C has engaged the NIH Tribal Health Research Office (THRO) to express our interest in making Tribal communities our partners, too.

Using Data Derived from Electronic Health Records to Study COVID-19

EHRs are digital, machine readable versions of patients' paper charts. They contain clinical information, such as medical history, diagnoses, demographics, immunization records, lab results, and medications. EHRs also contain data that may identify a person, also known as protected health information.

Under the [1996 Health Insurance Portability and Accountability Act (HIPAA)](#), covered entities — such as health care providers — may release data for research without obtaining an individual's authorization if direct identifying information is removed and appropriate oversight and agreements are in place. Accordingly, NCATS requested that health care providers transfer what is known as a "limited data set" to the N3C so that these needed clinical data could be used for COVID-19 research. A [limited data set](#) includes health information that removes 16 types of direct identifying information, which allows the data set to be used for purposes of research, public health, or health care operations without obtaining an individual's authorization; a data use agreement must describe the terms and conditions of the data use. NCATS requested the Limited Data Sets to include ZIP codes and dates of service, because these data are critical for tracking the progress of the pandemic. The data within the N3C involves minimal risk to individuals and NCATS received a waiver of consent by an NIH Institutional Review Board, conforming to the Federal Policy for the Protection of Human Subjects ("Common Rule").

NCATS establishes [Data Transfer Agreements](#) with health care organizations across the United States to provide these data to the N3C, and [Data Use Agreements](#) that specify the terms and conditions of data access by organizations that have researchers who wish to access the data. Once a health organization has signed the Data Transfer Agreement, arrangements are made for sending the Limited Data Set to NCATS. The participating health care providers update these data on a regular basis, approximately every 3-4 weeks, so that researchers can study the ongoing impact of COVID-19 on patients in the United States.

Data Protections

NCATS is balancing carefully the need for researchers' access to patient data to address the pandemic with the need to protect patient information and privacy. At the same time, NCATS and N3C encourage research collaborations and rapid dissemination of research results. NCATS works closely with its partners to discuss and develop ***regulatory and policy*** issues that affect the N3C operations; to develop ***privacy measures*** to keep the data protected; to test and monitor the Enclave and data cloud ***security***; and to review ***researcher responsibility*** policies to ensure that users do their part to keep the data secure and to protect the privacy of the patient health information. The following diagram shows the efforts NCATS has undertaken to ensure the security, confidentiality, and privacy of N3C data.

Regulatory & Policy	Privacy Measures	Security Testing & Monitoring	Researcher Responsibilities
• Data contributing sites abide by the HIPAA Privacy Rule • N3C activities fall under the Federal Policy for the Projection of Human Subjects in research ("Common Rule") • Data are provided as a HIPAA-defined Limited Data Set • NIH IRB oversight and waiver of consent • For COVID-19-related research only • No genomic data • No emergency public health authorities were used to obtain the data under these conditions	• Certificate of Confidentiality • Data stay within the Enclave; no download or capture of raw data • Privacy Impact Assessment • Review of project requests by Data Access Committee • Additional Tribal data privacy measures (while seeking a consultation with Tribal Nations)	• Federal Government-compliant Enclave managed by NCATS • Meets government security controls for cloud security and privacy • Data encryption in transit and at rest, without exception • Scheduled penetration testing • Active monitoring and logging by NIH and HHS • Auditing of activities in the N3C Enclave	• User's organization must have a Data Use Agreement with NCATS • Researchers' organizations also are responsible • Adherence to a Code of Conduct • Required NIH IT Security Training • Required Human Subjects' Protections Training • Follow Community Guiding Principles

Furthermore, contributing sites provide race and ethnicity data, including AI/AN, to the N3C as part of a data set collected from patient EHRs. Race and ethnicity information from health care settings is generally either self-identified or assigned by health care providers, does not relate to Tribal enrollment and does not include Tribal affiliation. For the AI/AN category, the data are aggregated under the "other" category. In addition, the N3C receives five-digit ZIP codes. Any ZIP code representing fewer than 20,000 individuals or that represents Tribal Land is obscured through aggregation.

Seeking Input from Tribal Leaders and Community Members

As part of our process, NCATS has an ongoing engagement with THRO, as well as other NIH experts who recently have completed a Tribal Consultation (e.g., RADx). NCATS is seeking Tribal Consultation for the N3C Enclave, where the origin of data is existing EHRs.

Since the establishment of the N3C, NCATS has obscured AI/AN data in the Enclave, while simultaneously working with the THRO to seek Tribal Consultation on whether and how to make N3C AI/AN data accessible to researchers. Until the Consultation is completed, AI/AN data will remain obscured. It is important to note that the N3C **does not** have EHR data from the Indian Health Service, nor does it contain information about Tribal affiliation. Finally, to ensure that researchers cannot use geographic information to infer potential Tribal affiliation, ZIP codes that overlap with Tribal communities are currently not available for studying specific research questions. Patient records where race and ethnicity data are self-identified as AI/AN are obscured temporarily as "other" to be available for total counts of patient numbers to reflect the accuracy of the numbers of patients within the N3C.

NCATS recognizes that the AI/AN data that are not currently accessible could be useful for understanding the impact of COVID-19, identifying effective interventions to improve public health, and generating knowledge directly useful to Tribal communities and public health

responses to COVID-19. Research conducted by accessing the N3C already is beginning to contribute to our understanding of COVID-19 and is revealing new knowledge, such as populations at greater risk for severe COVID-19 outcomes, COVID-19 severity when considering social determinants of health, identification of effective drugs repurposed as a treatment for COVID-19, and more. We would like the research utilizing the data in the N3C to inform the health of Tribal Nations, as well.

Ideally, NIH would have sought Tribal Consultation before the start of this program. However, given other COVID-related Consultations and urgency of the pandemic, NCATS decided to obscure AI/AN data until Consultation could occur. During the Consultation, NIH will seek input on whether and how to make AI/AN data available through N3C. If Tribal leaders are open to having more regular conversations, NCATS would welcome that opportunity.

Suggested Topics for Consideration

- Should AI/AN data be unobscured and made available for COVID-19 research?
  - How would Tribal communities want to be involved in governance, access, use, etc.?
- What steps would the N3C need to take to consider making the data available?
  - Which aspects of governance would be particularly critical for this resource?
  - How can ongoing partnerships with Tribal communities and researchers be instituted and maintained?
- If Tribal Nations identify benefits to using the N3C resource, what outreach is needed to better engage AI/AN researchers?