CTSA Steering Committee Meeting Summary
Zoom Conference
March 13, 2023; 2:30-3:30pm ET

Steering Committee Attendees:

Stephan Bour	Karen Johnston	Deborah Ossip	Lawrence Sinoway
Arleen Brown	Jessica Kahn	Ruth O'Hara	Randy Urban
Daniel Ford	Don McClain	Muredach Reilly	Rosalind Wright
Melissa Haendel	David McPherson	Doris Rubio	Ted Wun
Laura James	Duane Mitchell	Katheryn Sandberg	

SC Regrets:

Michael Kurilla, Teisha Johnson, Steven Reis, Martin Zand

Guest Speakers:

Emily Pfaff, UNC Chappell Hill School of Medicine

NCATS Attendees:

Audie Atienza	Sarah Dunsmore	Greg Jarosik	Joni Rutter
Jane Atkinson	Stephanie Ezequiel	Rebecca Katz	Clare Schmitt
Mathew Arnegard	Josh Fessel	Ken Gersing	Meredith Temple-O'Connor
Heather Baker	Stacia Fleisher	Carol Merchant	Salina Waddy
Patrick Brown	Gallya Gannot	Marilyn Moore-Hoon	Ken Wiley
Penny Burgoon	Brittany Gibbons	Anna Ramsey-Ewing	
Soju Chang			

Support Center:

CLIC: Rebecca Laird, Julie Schwan
CCOS: Kerry James, Lauren Fitzharris

Welcome & Review In-Person Agenda (Slides 1-4)

Speakers: Co-chair Duane Mitchell and Clare Schmitt

Dr. Mitchell and Dr. Schmitt welcomed the members of the Steering Committee members and facilitated the Steering Committee call. They reviewed the in-person meeting agenda at the end of the call.

2023 Spring and Fall Meeting Update (Slides 5-9)

Speaker: Kerry James, CCOS

Presentation summary: K James provided logistical details on the CTSA Spring Meeting including dates, time, and location (see below). The UL1 PI meeting on day 1 will be at a different hotel from the Steering Committee. Transportation will be provided between the two hotels to take CTSA members to the UL1 meeting and return them to the hotel hosting the Steering Committee meeting.

The registration numbers as of March 13, 2023 were: 345 total registrants, 49 for the UL 1 PI meeting, and 59 for the Steering Committee. One-third of registrations are virtual. Presentations will be streamed and the chat enabled during and directly after the presentation. Discussion groups and table discussions will not be streamed.

Registration for the spring meeting closes on Monday, April 3rd, 2023. There is a separate registration for the UL1 meeting and the Steering Committee and EC/Communicators meeting. A list of nearby food options for both hotels will be made available in advance of the meeting.

2023 CTSA Spring Meeting Details:

- UL1 PI Meeting: Thurs, 4/20 from 11am-noon at the Capital Hilton
  - Registration: https://web.cvent.com/event/cc7f9920-e3d4-4b8a-b17c-b691a0e5558a/regProcessStep1
- Steering Committee Meeting: Thurs, 4/20 from 1:30-6pm at the Washington Hilton
  - Registration: https://cvent.me/aLEzny
- Enterprise Committees and Hub Communicators: Fri, 4/21 from 9am-4pm
  - Registration: https://cvent.me/aLEzny
- Spring Meeting Webpage: 2023 CTSA Program Spring Group Meetings | clic (clic-ctsa.org)

Registration as of Date	**2023-03-13**
Total Registration	**345**
Total Virtual	**121**
Total In-Person	**224**

CTSA Program Meeting	In-Person	Virtual	Total
April 20, 2023; 11AM-12NOON			
UL1 Principal Investigator	**49**	**N/A**	**49**
April 20, 2023; 1:30-6PM			
Steering Committee	**50**	**9**	**59**
April 21, 2023; 9AM-12NOON			
Collaboration & Engagement EC	**62**	**35**	**97**
Integration Across the Lifespan EC	**22**	**16**	**38**
TRIAD	**96**	**45**	**141**
April 21, 2023; 1-4PM			
Communicators Group	**43**	**25**	**68**
DEIA EC	**74**	**38**	**112**
Informatics EC	**43**	**21**	**64**

K James gave up-to-date details on the 2023 Fall Program Meeting. Tentative dates are November 6-8, 2023. The contract for the Double Tree Hilton (Crystal City) is in progress. It will be a 3-day meeting, with a post session likely hosted in the evening on day 2.

N3C CTSA site data score cards: understanding site and consortia data variability (slides 1-13)

Speaker: Emily Pfaff

Presentation summary: Dr. Pfaff gave an overview of the N3C scorecard and started with a review of the established workflow of N3C: Data are submitted in one of four common data models (OMOP, TriNetX, PCORnet, or i2b2/ACT) securely to N3C. Data are pulled from what is called the ingestion and harmonization pipeline and the data are transformed to comply with OMOP. COVID and non-COVID related quality checks are completed against the data.

As N3C began to grow, the N3C team realized the data should be returned to the site so that each CTSA hub has the opportunity to make repairs and changes to local data. This is the point where the scorecard initiative was created. The main goals of the project are to 1) inform all data partners about the quality of their data; 2) stop quality regression and maintain data quality, and 3) benchmarking to compare anonymously across sites.

Dr. Pfaff showed examples of interpreting the scorecard, including using basic visualizations and heatmaps to provide important information. When sites send lab and measurement data that have null units of measure or invalid units, N3C now has an inference pipeline built in to add units or fix them when they're invalid. This "rescued information" is sent back to sites so fixes can be implemented in the local environment. In lieu of reading 60 pages of data quality information, N3C is developing an executive summary called the PI Scorecard which will be available in the next few weeks.

Launching the National "Clinical" Cohort Collaborative (N3C) (slides 14-30)

Speaker: Melissa Haendel

Presentation summary: Dr. Haendel summarized impact stories, including how N3C won the grand prize in the Dataworks Competition. N3C has had diverse impact by transforming a number of different care guidelines, developing evidence-based disease definitions, and having a variety of complex risk prediction models.

In just two years, N3C has had over 1900 citations and an H index of 24. Additional evaluation measures have found 3,400 users and over 430 studies for team science. N3C is primarily led by women in minority leadership, and the N3C data set is currently the largest limited data set publicly available in the United States with over 18M patients.

In addressing bias, Dr. Haendel noted that N3C is representative demographically in terms of race, age, ethnicity, sex, rurality, geographic location, and socioeconomically. Existing biases including data sites that are academic medical centers—primary care centers with sophisticated informatics teams. Additionally, clinical encounters and real-world data from EHRs are not representative of a person's health, only their clinical encounters. Integrating harmonizing data across medical centers enhances power, reduces biases, and enables new opportunities like rare disease research.

Similarities between the pilot program called N3C Clinical and N3C are they are 1) continuing to leverage common data model (CDM) repositories such as OMOP, PCORNet, TriNetX, 2) continuing to provide executable queries for a specific CDM that will extract patients with a longitudinal connection to a clinical organization, 3) continuing centralized harmonization and data quality enhancements such as imputing missing units of measure, 4) continuing to provide a highly secure, FISMA-moderate compliant, data analytic environment that blocks any data exfiltration, 5) continuing to operate on a federally managed cloud.

The differences between this pilot and the original N3C is that contributing organizations have complete agency and control over how their data is used. In the pilot, each institution can choose who has access to their data.

Community Governance calls are on Fridays at 7pm ET. The meeting link can be found on the N3C calendar: https://covid.cd2h.org/n3c-calendar

Questions and Discussion (for Dr. Pfaff and Dr. Haendel's presentations):

- The N3C Clinical Pilot will include the scope of EHR data that is being submitted for COVID, but there are several external and linkable data sets within N3C that increase the scope without additional efforts from sites. The pilot phase is not at the point where data, such as imaging, will be added.
- When comparing N3C with the PCORNet approach, the data brought together centrally have revealed differences and data quality issues that mean the data are not truly comparable and consistent. There are a number of different models that N3C may consider longer term to help distributed networks more effectively from what N3C has learned by centralization.
- There are certain checks that must be met before data are released. For example, if data are submitted without COVID tests, N3C will not release the data. There are "heads up checks" which indicate that data are a couple standard deviations away from the mean and a site may want to examine their data but does not prevent the data from being released. It's up to some sites to decide what to do with the information, and some sites may not have the capacity to make changes.
- A presentation and summary data of the different pilots and projects that are underway (e.g., N3C pilot, NCI pilot, N3C COVID) would be helpful.
- There is not a lot of time at the in-person meeting to have a prolonged discussion on N3C. The Steering Committee will take into consideration follow-up from this robust discussion on how to engage both the Steering Committee and PODs for additional feedback and discussion.

Meeting Action items

Date	Action	Person Responsible	Status
	None for this meeting		

Next Steering Committee Meeting: Monday, March 27, 2023 at 2:30pm ET