It Worked for “All of Us”

Unlocking Data for Research

Ray Duncan, MD, Executive Director
Technology and R&D, Enterprise Information Services
Cedars-Sinai Health System

Erica Galvez, CEO
Manifest MedEx

Josh Lemieux, HIN/HIE Lead for All of Us Research Program, Center for Linkage and Acquisition of Data
University of North Carolina at Chapel Hill

Jay Nakashima, President
eHealth Exchange

Speakers

Josh Lemieux
University of North Carolina at Chapel Hill

Erica Galvez
Manifest MedEx

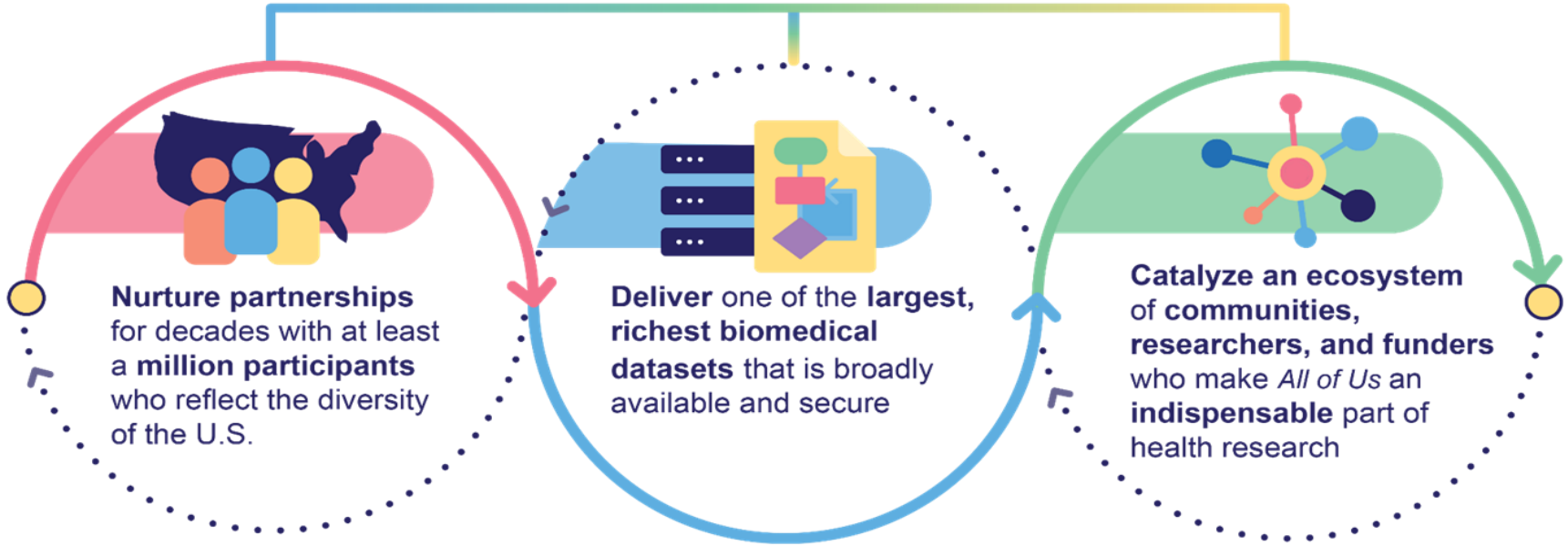
Ray Duncan, MD
Cedars-Sinai Health System

Jay Nakashima
eHealth Exchange

NIH *All of Us* Research Program

All of Us Research Program

Mission: Accelerate health research and medical breakthroughs, enabling individualized prevention, treatment, and care for all of us.



Made possible by a team that maintains a culture built around the program's core values

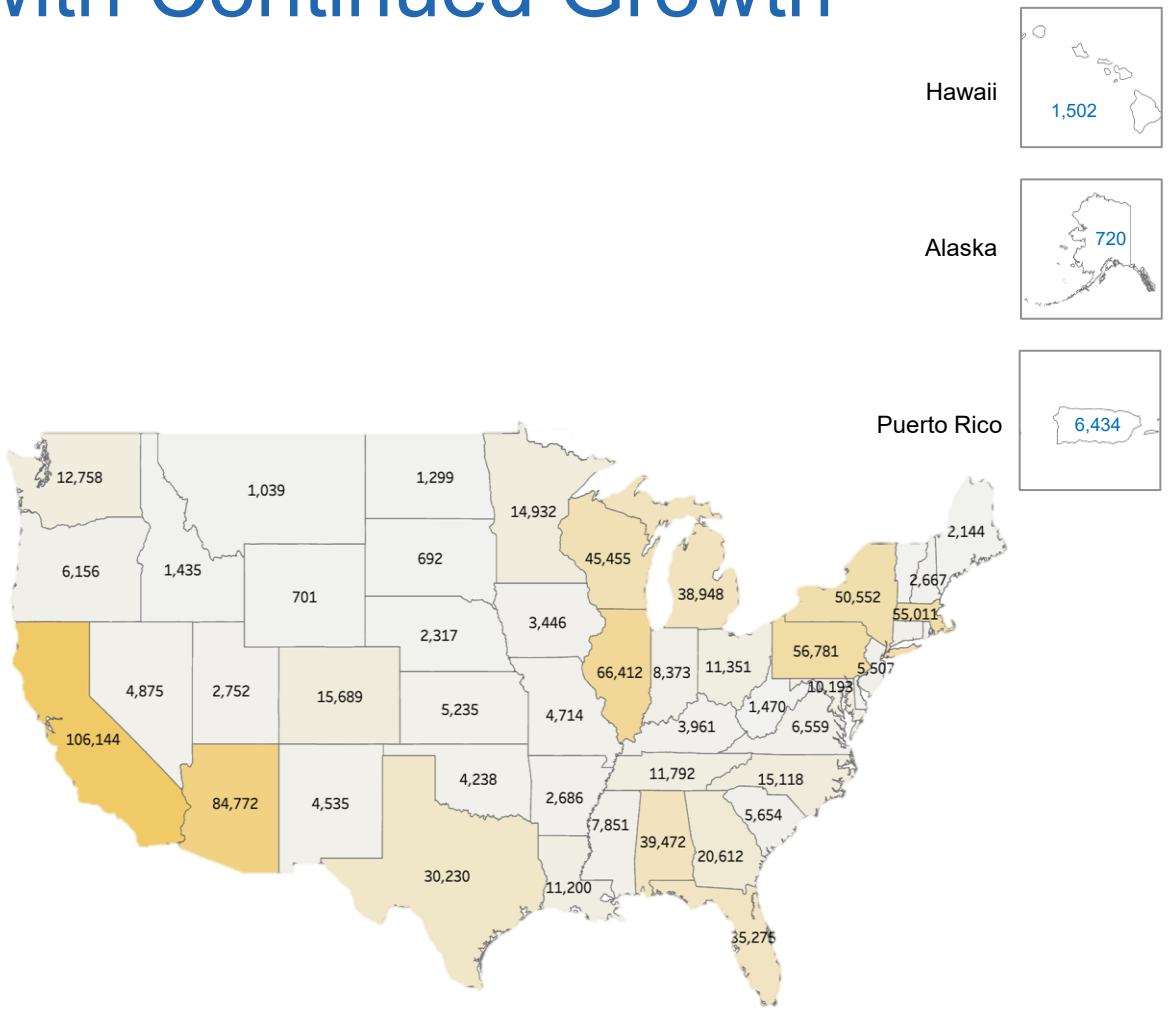
All of Us Research Program with Continued Growth

849,000+
Participants

454,000+
Electronic Health Records

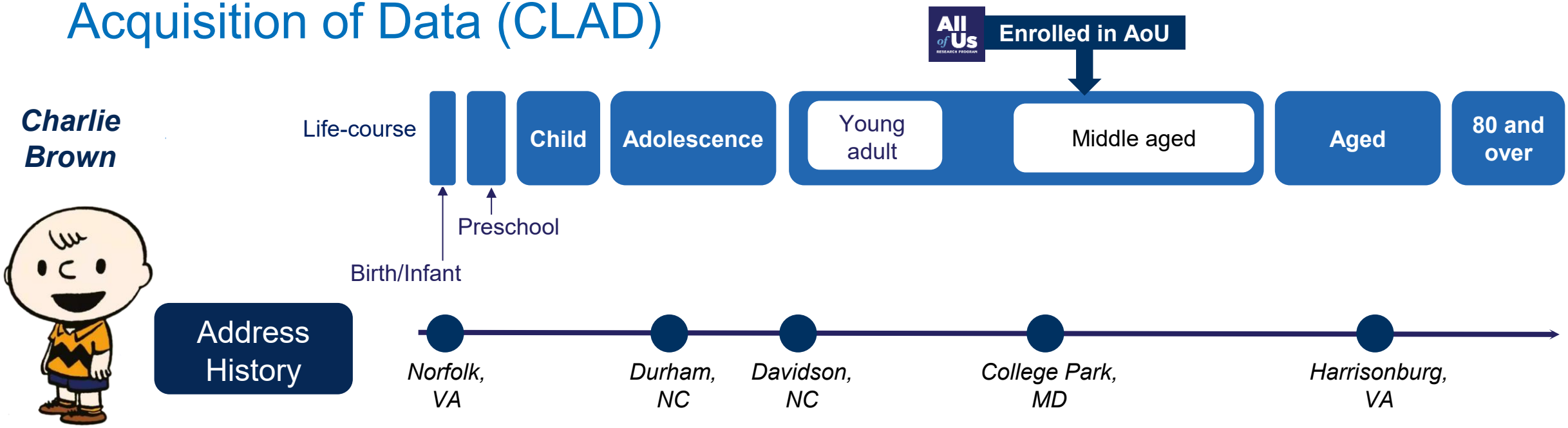
574,000+
Participants who have completed the initial steps of the program

597,000+
Biosamples



Numbers current as of November 25, 2024

Challenge Addressed by *All of Us* Center for Linkage and Acquisition of Data (CLAD)

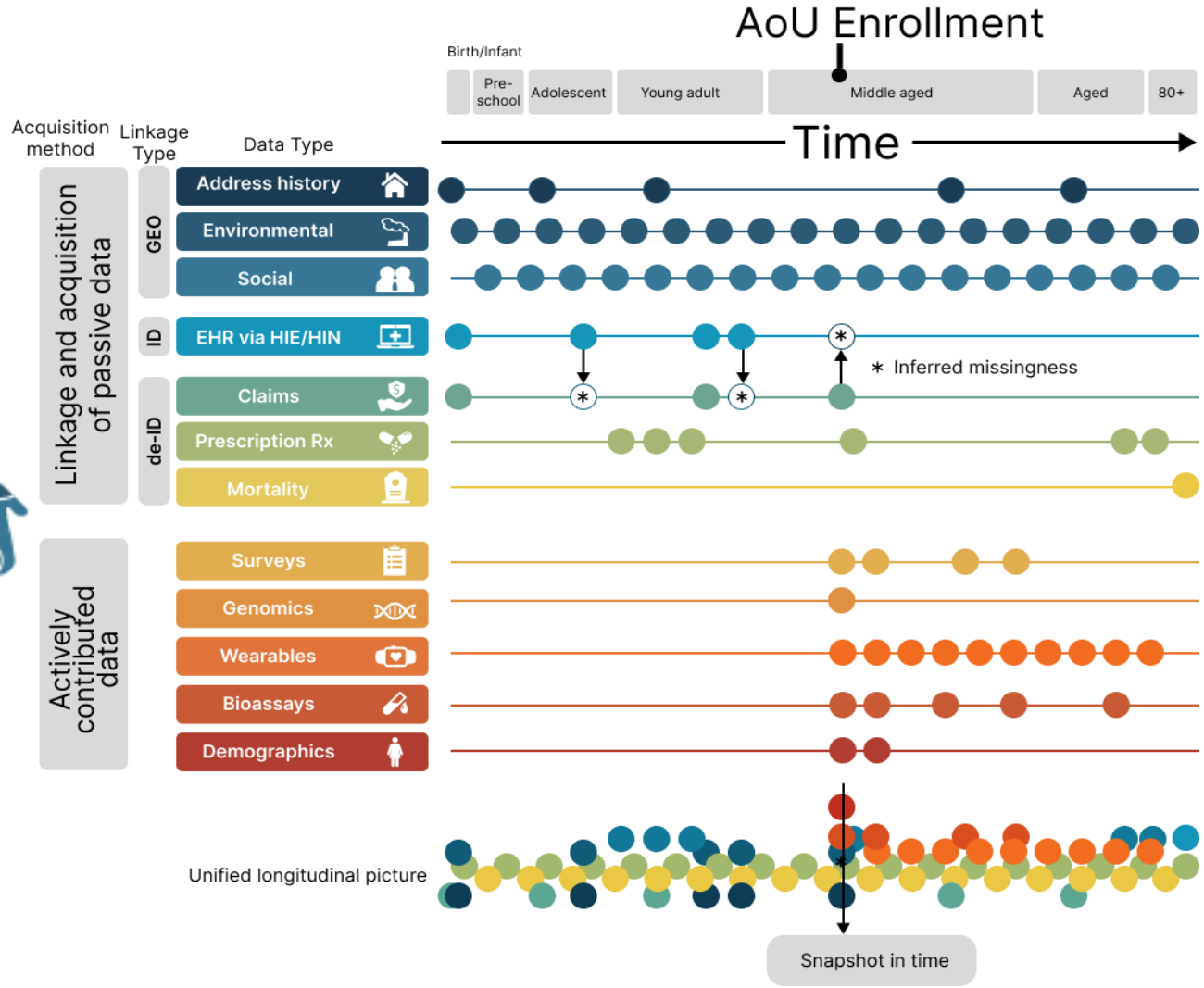


- Data about individuals is scattered across health systems and numerous sources
- A comprehensive dataset for each individual's life-course could improve health and research
- Many hurdles: different formats and standards, regulatory and legal, security and technical compatibility

CLAD Vision: Putting the Patient Back Together Again

New data acquisition methods are needed to acquire and link data to each person:

- Patient-Privacy Preserving Record Linkage (PPRL, de-identified token-based linkage)
- Geo-spatial & temporal linkage
- HIE/HIN identified linkage to acquire EHR data for research

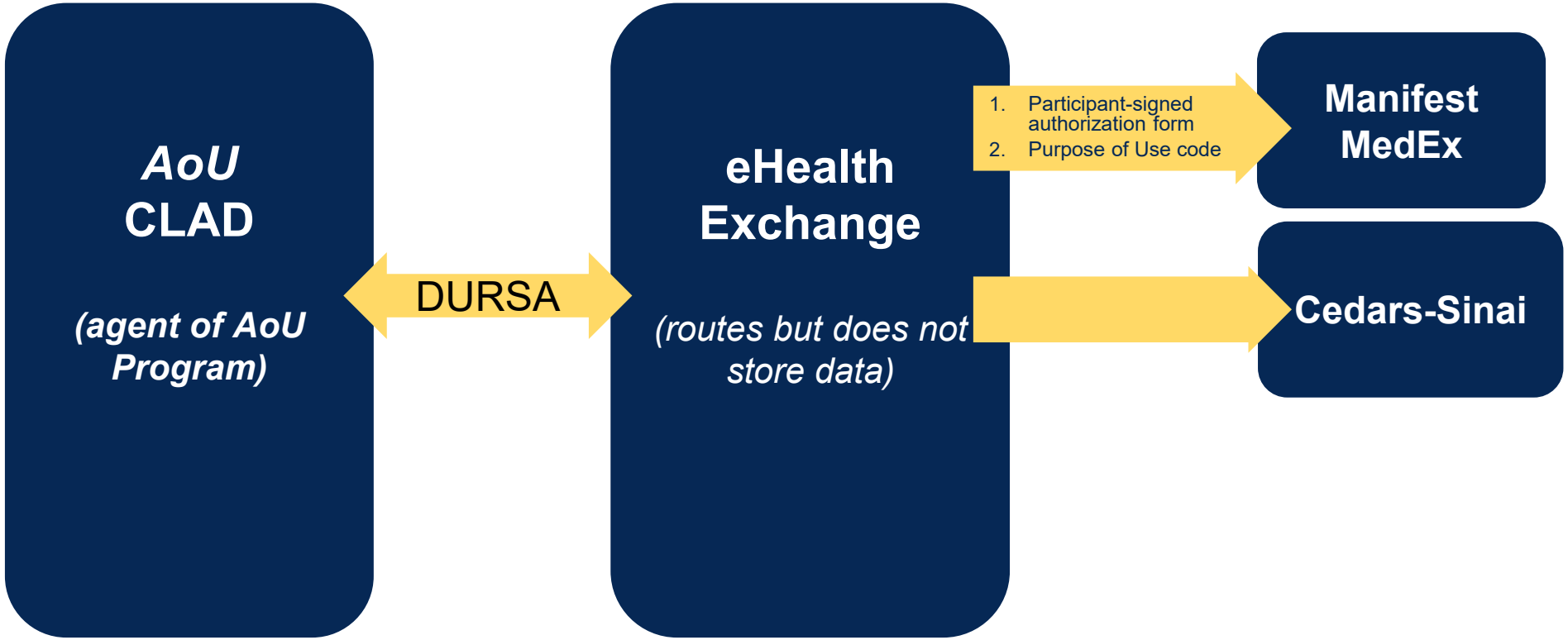


All of Us Research Program Pilot

Trust Framework

Trust Approach:
eHealth Exchange
Data Use and Reciprocal Support Agreement (DURSA)

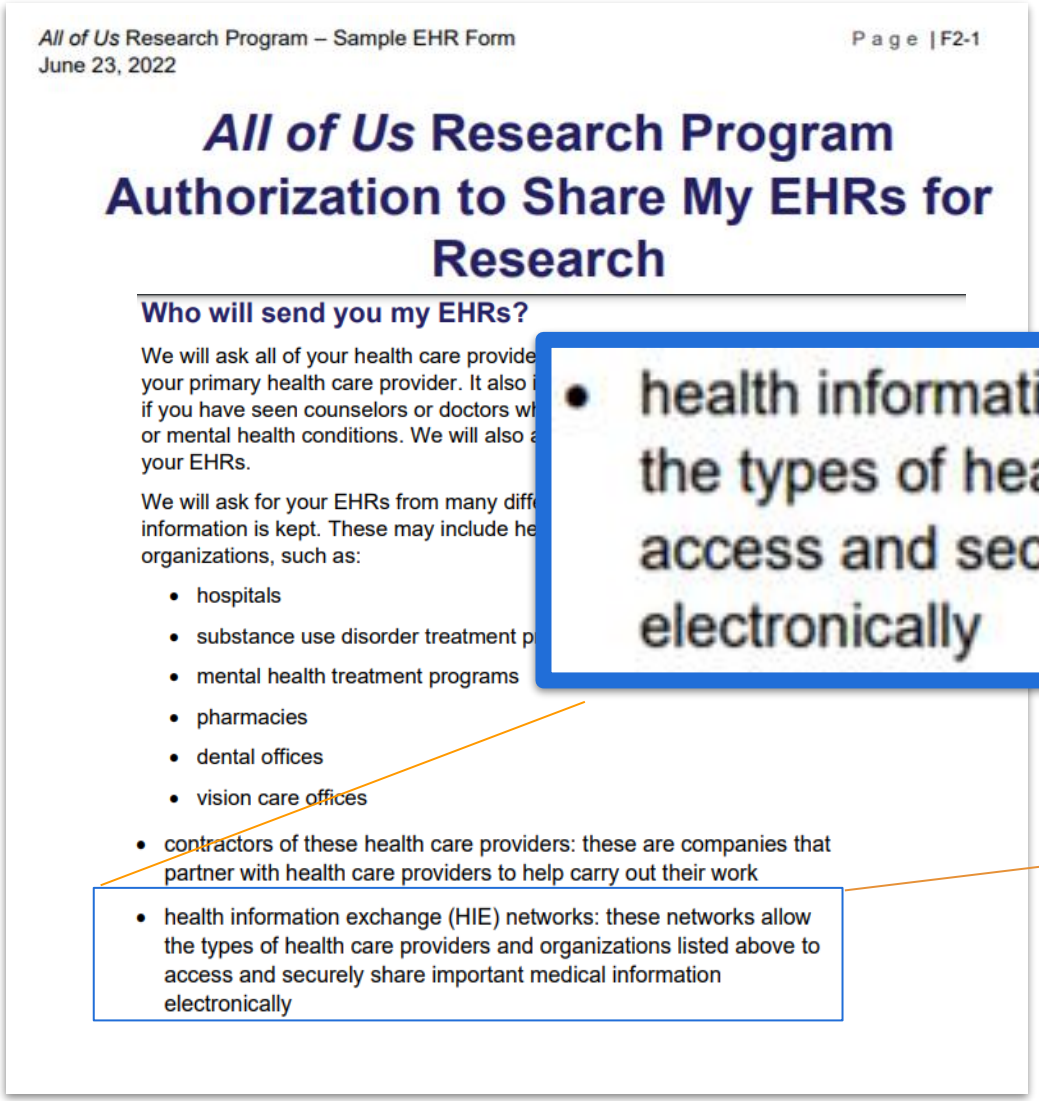
Signed by all 325 network members



Under DURSA, the pilot conducts *AoU* participant authorization-based queries as a permitted use of eHealth Exchange network

*For pilot, Purpose of Use Code will be "RESEARCH" (see Page 23 and 24, Table 3.3.2.6-1, of Nationwide Health Information Network Authorization Framework Specification V 3.0)

Authorization-Based Exchange for Research



All of Us Research Program – Sample EHR Form
June 23, 2022

Page | F2-1

All of Us Research Program Authorization to Share My EHRs for Research

Who will send you my EHRs?

We will ask all of your health care provide
your primary health care provider. It also i
if you have seen counselors or doctors wh
or mental health conditions. We will also a
your EHRs.

We will ask for your EHRs from many diff
information is kept. These may include he
organizations, such as:

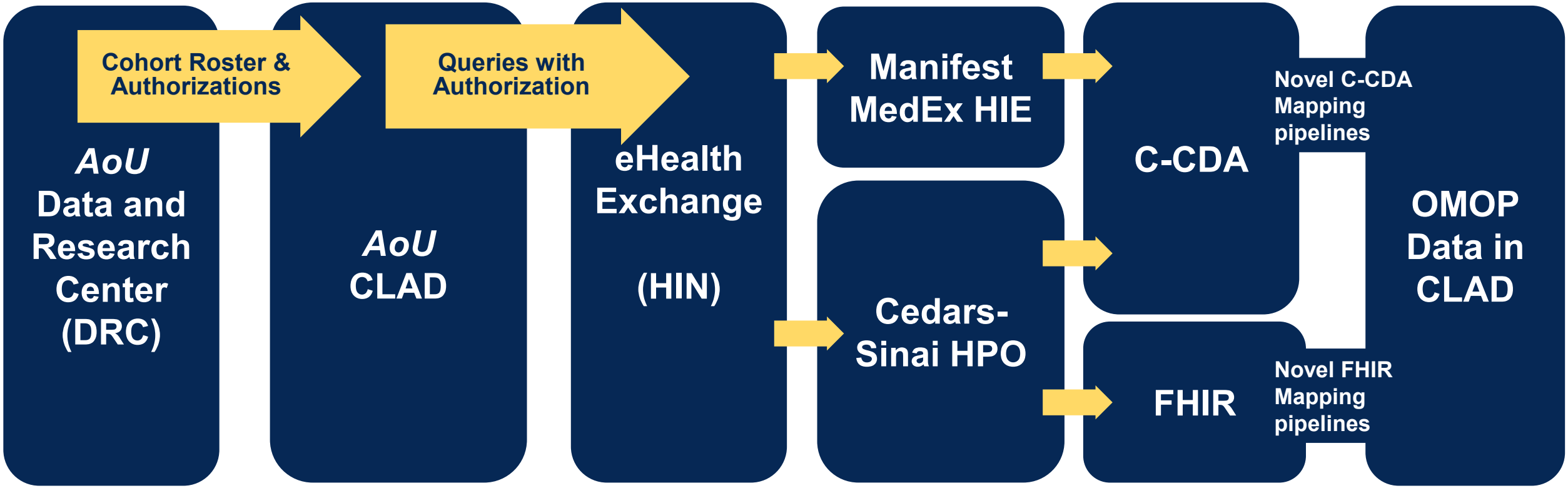
- hospitals
- substance use disorder treatment p
- mental health treatment programs
- pharmacies
- dental offices
- vision care offices

- contractors of these health care providers: these are companies that partner with health care providers to help carry out their work
- health information exchange (HIE) networks: these networks allow the types of health care providers and organizations listed above to access and securely share important medical information electronically

- health information exchange (HIE) networks: these networks allow the types of health care providers and organizations listed above to access and securely share important medical information electronically

Using HIE/HINs EHR Data for Research: A Pilot

~94% of AoU participants consent to share EHR data, yet only **~69%** have records in AoU

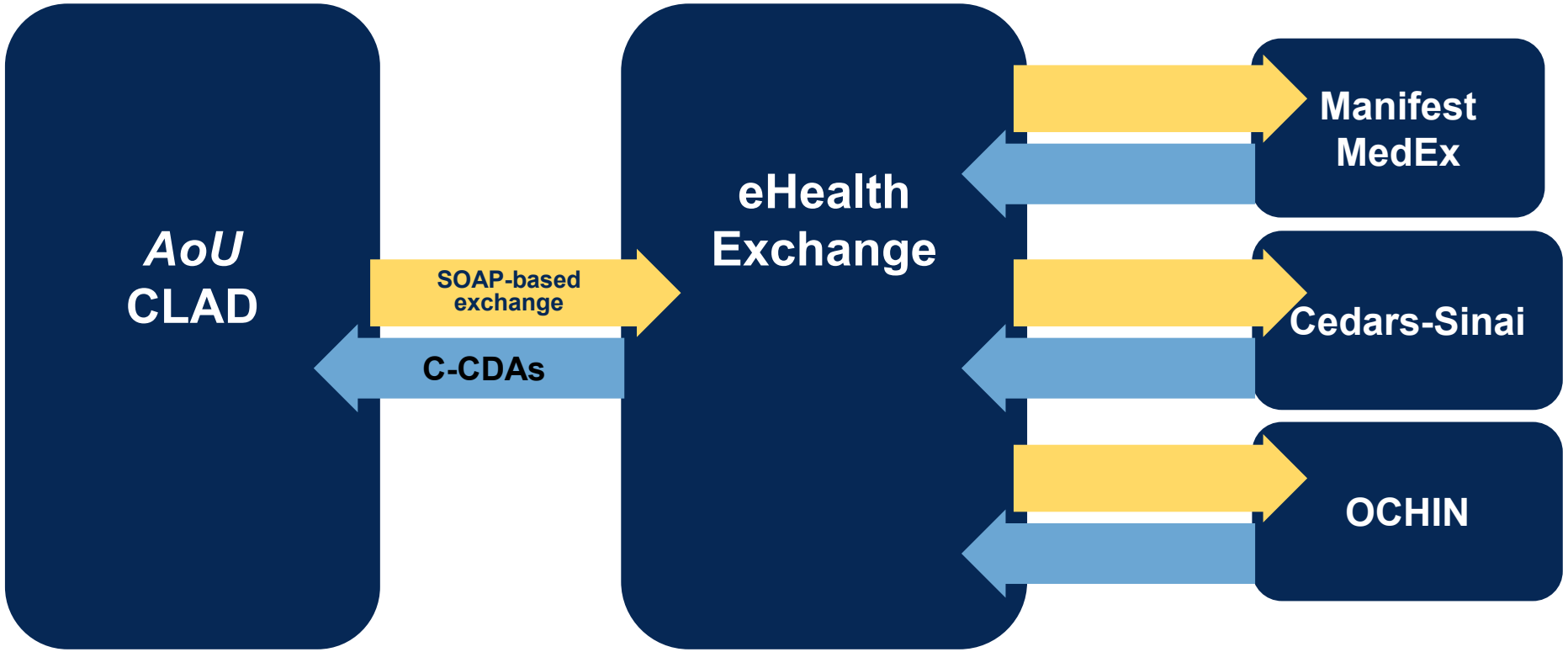


First success at this scale acquiring EHR data from national HIN/HIE networks pursuant to authorizations for research

CLAD Pilot Data - Aim 1: SOAP-based, C-CDA Payload

AIM 1 Technical Approach: Demonstrate the technical feasibility of obtaining EHR-sourced AoU participant standardized data on a convenience sample of AoU participants based on region(s) of pilot HIE(s)

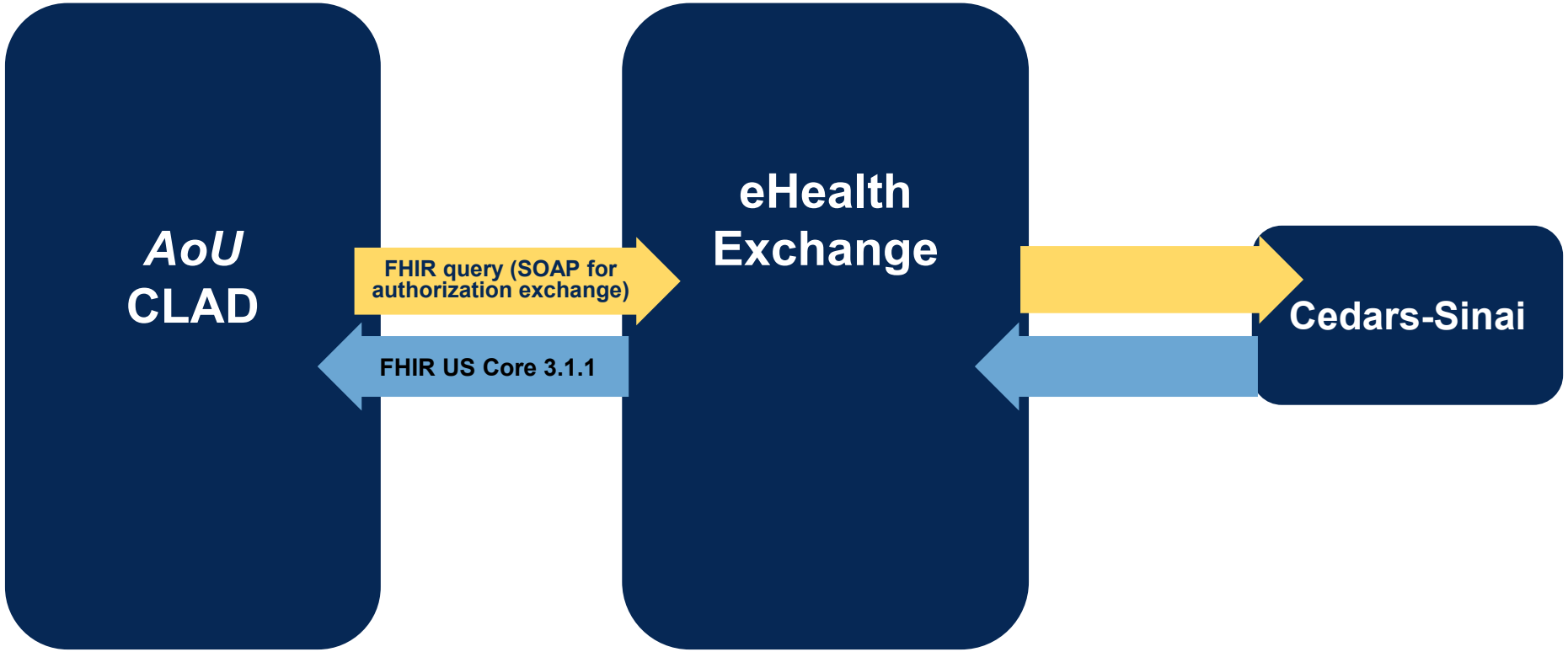
Patterned after Social Security Administration (SSA) model for exchange authorization PDFs.



- The transaction standard is **Simple Object Access Protocol (SOAP)**
- The standardized data payload will be **C-CDA/ CCD (2.1) compatible with 21st Century Cures Act**.

CLAD Pilot Data - Aim 2: FHIR

AIM 2 Technical Approach: Demonstrate the technical feasibility of obtaining EHR-sourced AoU participant standardized data on the same convenience sample of AoU participants as Aim 1, querying for records spanning the same time period as Aim 1 queries.

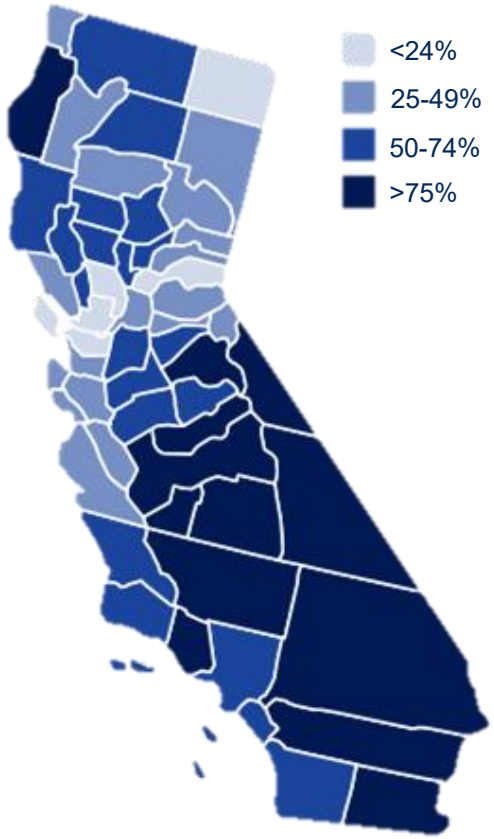


- This aim uses the same SSA model-based trust mechanism as Aim 1 but will to test **HL7® FHIR® (Fast Healthcare Interoperability Resources)** as the message transaction standard
- The data payload in **USCDI V2 as operationalized by FHIR US Core 3.1.1.**

All of Us Research Program Outcomes
Manifest MedEx

About Manifest MedEx

- Mission: provide every healthcare organization in California with the information they need to improve care, enhance health, and lower costs
- Largest nonprofit health information organization in California
  - Providing claims, clinical, and lab data for 50M individuals across 19 health plans, 140+ hospitals, and 2600+ providers.
  - 3M+ admission, discharge, and transfer (ADT) notifications delivered monthly
- CalHHS Data Exchange Framework (DxF) Qualified Health Information Organization (QHIO)
- Interoperability with 70+ EHRs
- Access to national networks via eHealth Exchange, TEFCA
- HITRUST CSF Certified since 2019
- NCQA Validated Data Streams since 2021



HITRUST

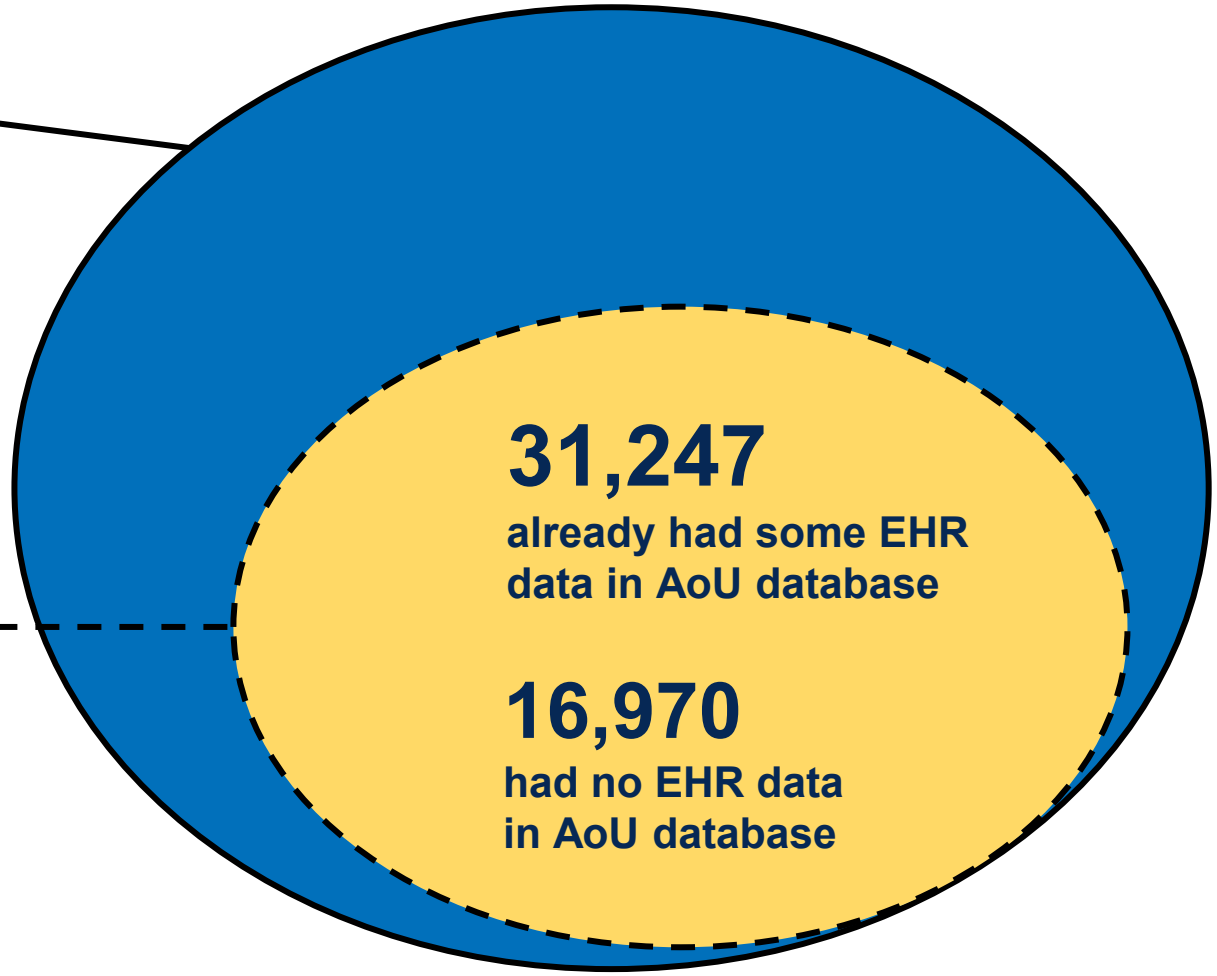
DATA EXCHANGE FRAMEWORK QUALIFIED HIO

Very High Match Rate with Statewide HIE

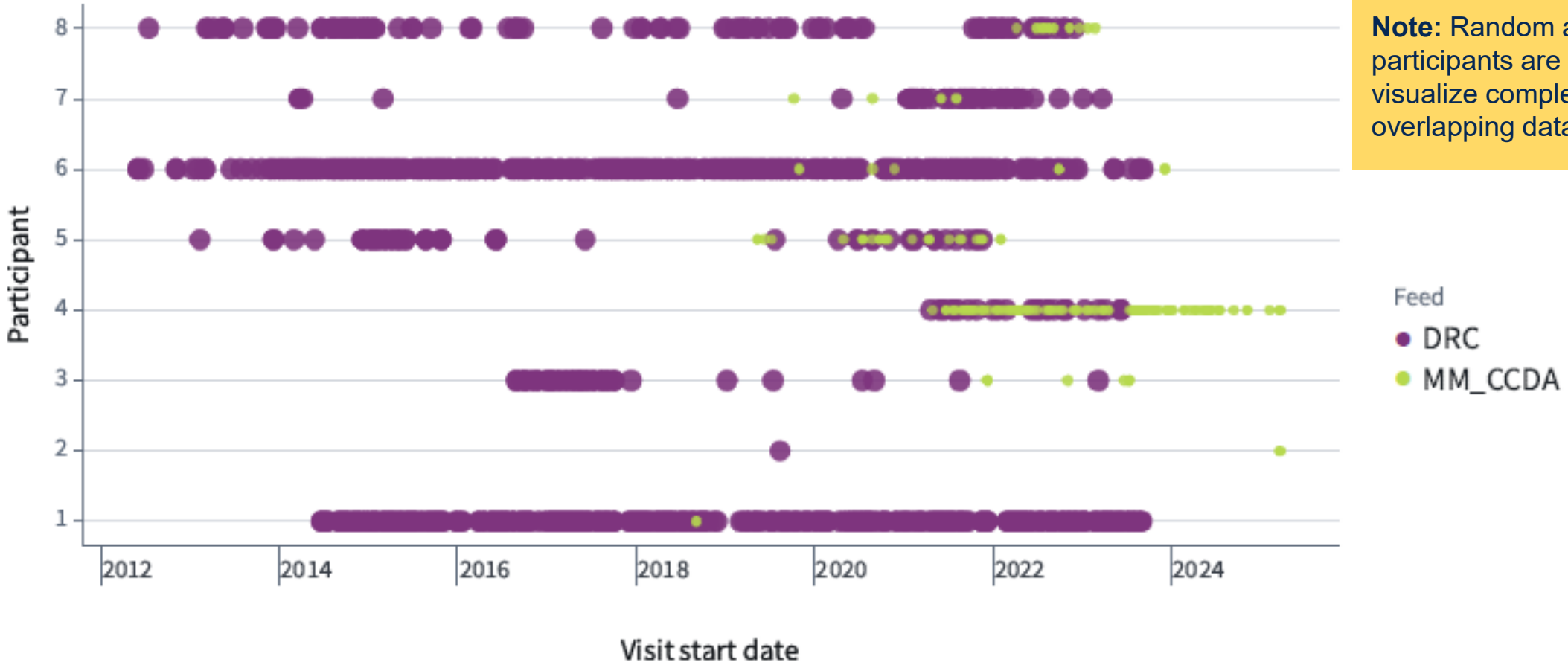
78,020
Manifest MedEx Queries

>60% match

48,217
Manifest MedEx C-CDA data received



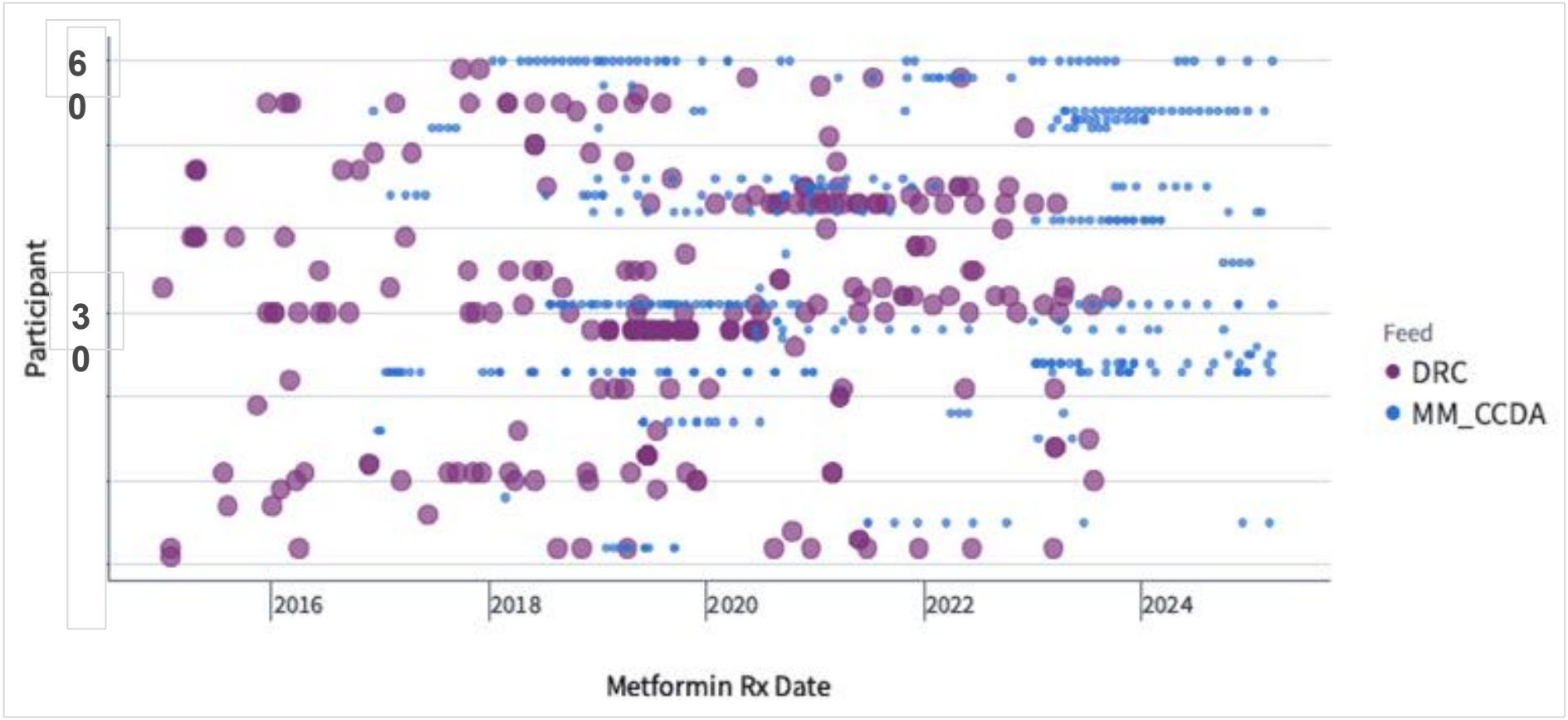
Manifest MedEx C-CDA vs DRC: Visit Timeline for 8 Random Patients



Note: Random anonymized participants are used to visualize complex patterns of overlapping data

The Manifest MedEx data adds additional visits not captured in the DRC data, mostly in and after 2020.

Manifest MedEx vs DRC: Drug X Prescription Records for 60 Random Patients

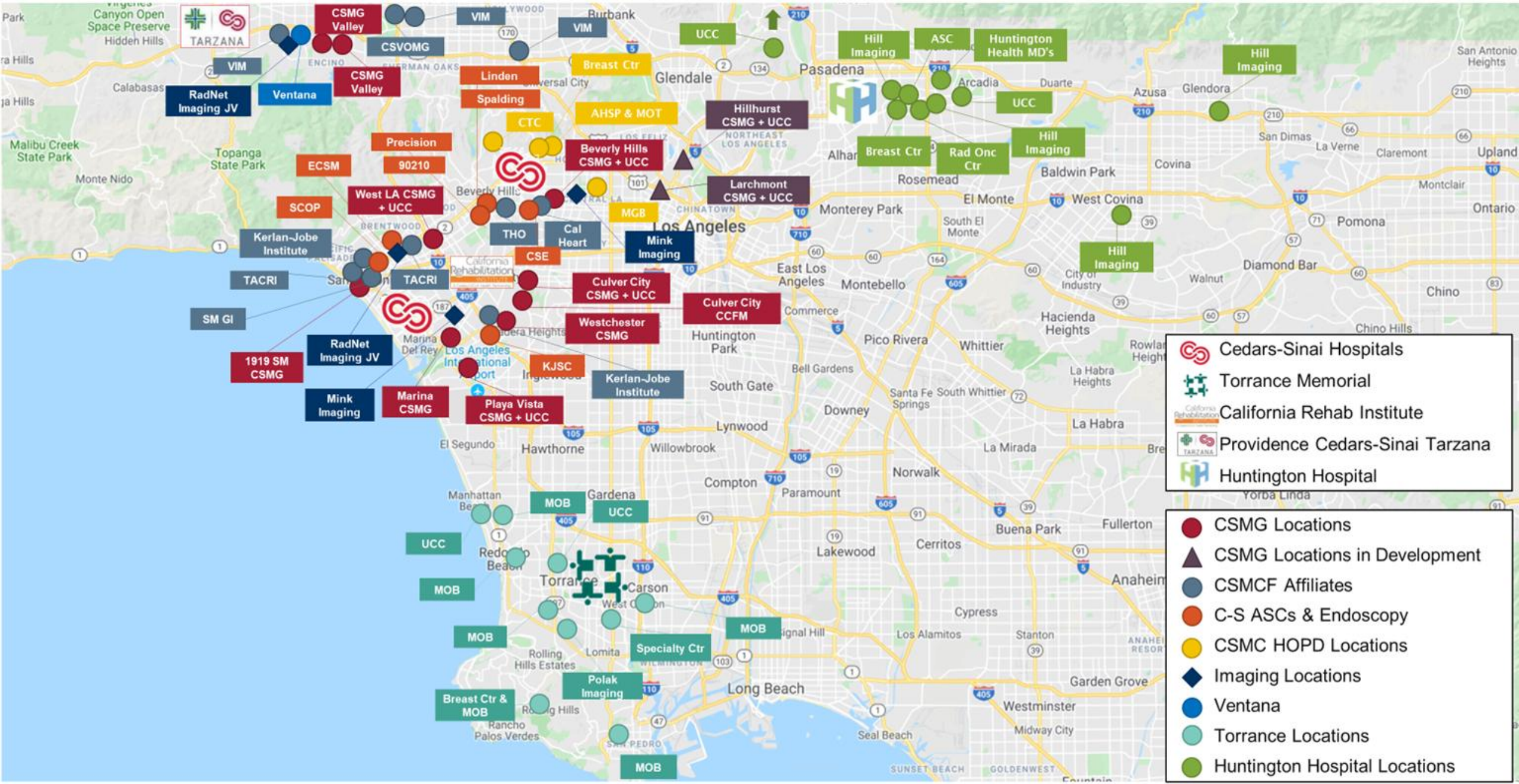


CCDA data from Manifest Medex (blue dots) adds many additional Drug X records to DRC data (purple dots) for these 60 randomly selected patients taking Drug X.

All of Us Research Program Outcomes
Cedars-Sinai

Cedars-Sinai Health System

- Established 1902
- Main hospital 950 licensed beds, 150 adult and peds ICU beds, 45 bed NICU, Level 1 Trauma Center
- 3 smaller community hospitals 150-500 beds ea.
- Rapidly expanding outpatient footprint: urgent care centers, outpatient clinics, and specialty affiliates
- 15,000 + employees, 3,096 Physician medical staff, 450 residents and fellows, 3,500+ nurses
- Epic Systems HIS/EMR fully deployed (2008)
- HIMSS Level 7 (2014), Most Wired Level 10 (2024)
- US News #1 in California 2025 (tied with UCLA)
- Separate Biomedical Computing Department
- Medical Informatics Fellowship
- Research Institute
- Comprehensive Cancer Center
- Comprehensive Transplant Center
- Heart Institute
- Neurosurgical Institute
- California Rehabilitation Institute (joint venture with UCLA)
- Tarzana Hospital (joint venture with Providence)
- Active community outreach programs (e.g. COACH for Kids)


TARZANA
CSMG Valley
CSVOMG
VIM
VIM
VIM
RadNet Imaging JV
Ventana
CSMG Valley
Linden
Spalding
Breast Ctr
AHSP & MOT
CTC
Precision
90210
ECSM
Beverly Hills CSMG + UCC
West LA CSMG + UCC
SCOP
MGB
THO
Cal Heart
Mink Imaging
CSE
Kerlan-Jobe Institute
TACRI
TACRI
Culver City CSMG + UCC
Culver City CCFM
Westchester CSMG
SM GI
RadNet Imaging JV
1919 SM CSMG
KJSC
Kerlan-Jobe Institute
Mink Imaging
Marina CSMG
Playa Vista CSMG + UCC
UCC
Hillhurst CSMG + UCC
Larchmont CSMG + UCC
Hill Imaging
ASC
Huntington Health MD's
Hill Imaging
UCC
Breast Ctr
Rad Onc Ctr
Hill Imaging
Hill Imaging
Hill Imaging
UCC
MOB
Gardena
MOB
MOB
MOB
Specialty Ctr
Polak Imaging
Breast Ctr & MOB
MOB
Cedars-Sinai Hospitals
Torrance Memorial
California Rehab Institute
Providence Cedars-Sinai Tarzana
Huntington Hospital
CSMG Locations
CSMG Locations in Development
CSMCF Affiliates
C-S ASCs & Endoscopy
CSMC HOPD Locations
Imaging Locations
Ventana
Torrance Locations
Huntington Hospital Locations

Cedars-Sinai and the *All of Us* Research Program

- Cedars-Sinai started enrolling patients in *All of Us* in June 2018
- Funded by a 5-year NIH grant to the "California Precision Medicine Consortium," Co-PIs Dr. Spencer SooHoo (EIS) and Dr. Mark Goodman (SOCC)
- Grant renewed in 2024 for another 5 years
- Approximately 8500 patients enrolled in *All of Us* at Cedars-Sinai to date, patient consents are held at the AoU DRC
- EHR data extracts are sent to the DRC quarterly. Data are de-identified by the DRC and loaded into the *All of Us* Researcher Workbench
- Patients can also enroll online directly.
- The *All of Us* "Sync for Science" mobile app was able to connect to the EHR and contribute data to the project directly (released in 2019, withdrawn in 2024)

Match Rates with Cedars-Sinai – an LA Area HPO

Cedars Sinai

Cedars-Sinai is an AoU HPO based in LA Area

Patient Discovery*

Total Queried	No Match Found	Participants Matched
78,020	70,851	7,173

Participant Match Rate*: 9.2%

CCDA Document Discovery

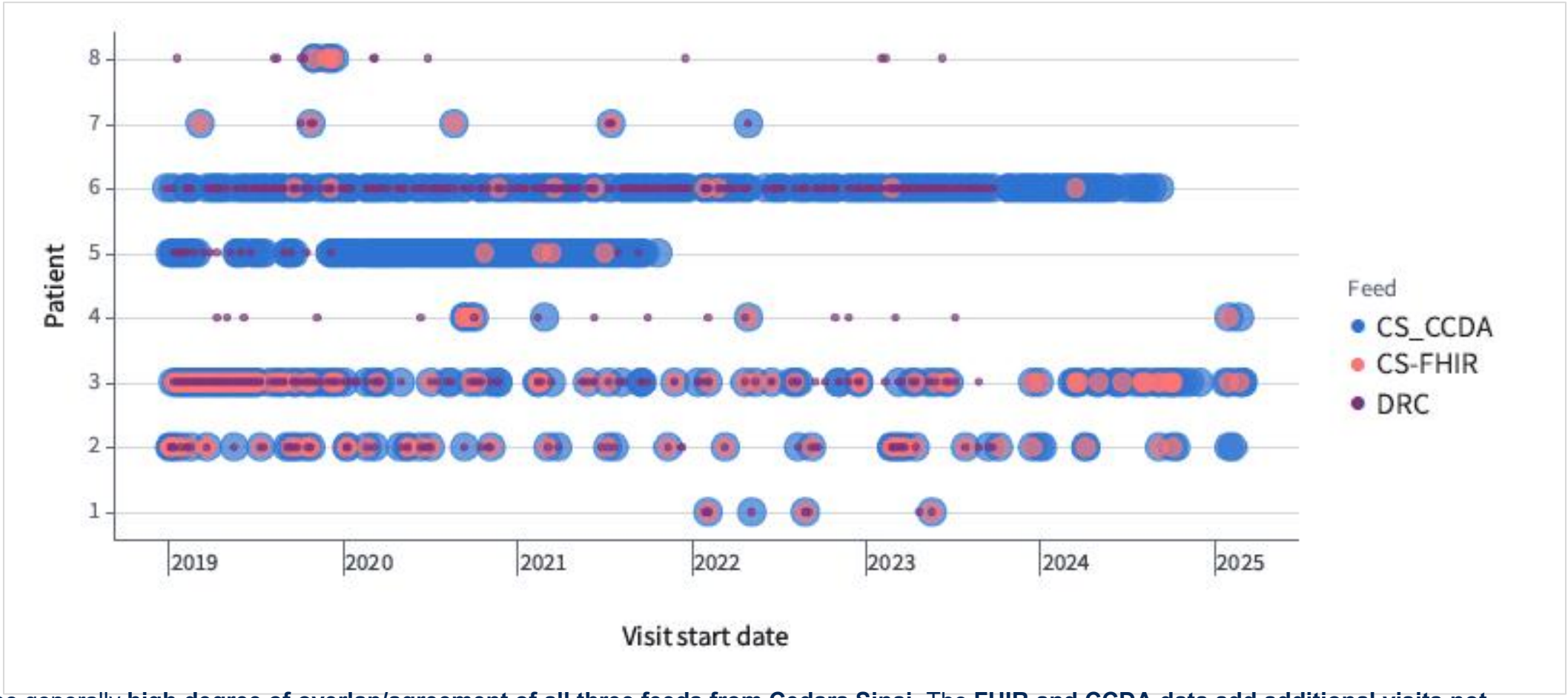
Documents Retrieved
7,156

FHIR Resource Discovery

Participants Retrieved	Resources Retrieved	Resources / Participant
7,119	9,872,388	1,387

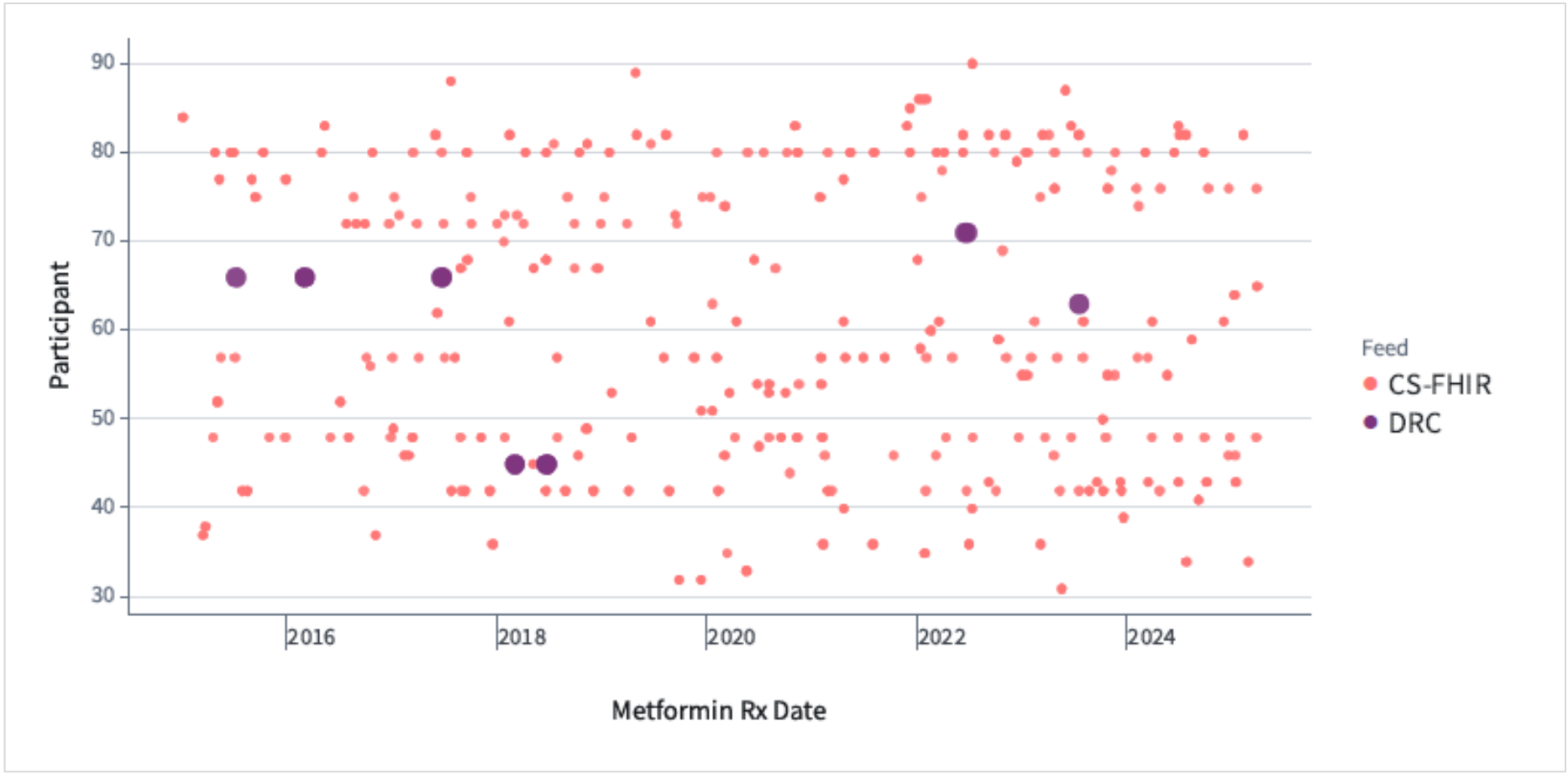
Context: Cedars-Sinai lower match rate than Manifest MedEx was expected. Both received roughly 78,000 queries for all California participants. Unlike Manifest MedEx, Cedars-Sinai is not a statewide HIE and therefore is not expected to have as many matches. However, duplicate records may exist in Manifest MedEx due to HIE via CeQ.

Cedars Sinai FHIR vs Cedar Sinai CCDA vs DRC: Visit Timeline for 8 Random Patients



Note the generally **high degree of overlap/agreement of all three feeds from Cedars Sinai**. The **FHIR and CCDA data add additional visits not captured in the DRC data**, mostly in 2023 or later. The exceptions are worth noting. Patient 5 does have a large number of pre-2022 visits not captured in the DRC data, but captured in the CCDA and FHIR feeds. Conversely, patients 4 and 8 have many DRC visits not captured in FHIR or CCDA.

Cedars Sinai FHIR vs DRC: Drug X Prescription Records for 60 Random Patients



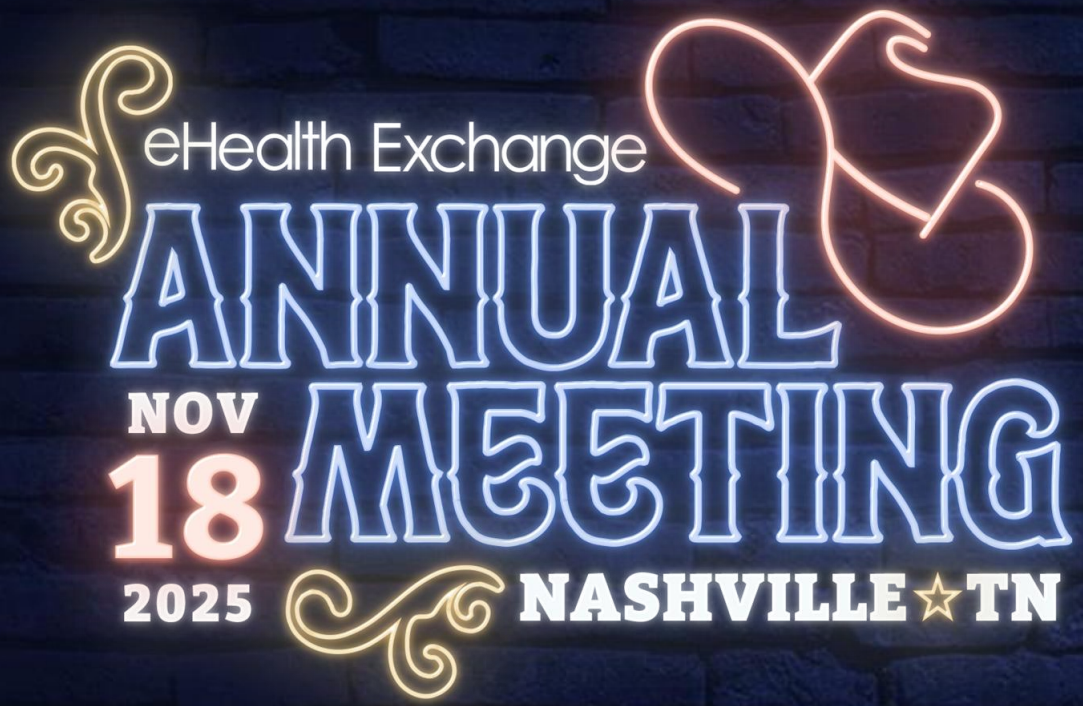
Cedars Sinai FHIR (pink dots) adds significant numbers of Drug X prescription records to the data already held by the DRC (purple dots). This indicates potential of FHIR to access specific FHIR Resources (data domains) that might be missing.

Some Challenges and Lessons Learned

- C-CDA interface used by SSA to retrieve charts and store authorization in the EHR could only be used for treatment purpose of use. Separate interface had to be licensed for research purpose of use. Installation was delayed by a scheduled CSHS Epic upgrade.
- Licensed C-CDA interface did not function as expected. Patient consents were retrieved from DRC via an S3 bucket and bulk-loaded in CS EHR.
- Contractual and technical delays pushed back end-to-end testing.
- Break the Glass protection on some patients initially caused failures of C-CDA and FHIR retrievals, requiring security fix to background service account.
- Bulk conversions of EHR data from a recent affiliation had diverse encounter types and other issues, making data difficult to map for analysis purposes.
- Discrepancies between bulk extracts, C-CDAs, and FHIR for encounter data, medication data, and immunization data have diverse causes that need investigation and follow-up.

Conclusions

- Consider research requests/responses for consented research projects under DURSA
- Encourage your local providers to change their HIE settings to respond to research-purposed queries authorized by the patient
- Keep eyes on TEFCA; this projects shows potential of yet another benefit of fully networked nationwide exchange
- Data quality infrastructure/continuous repair and feedback loops will be critical to improve health through analysis of longitudinal real-world data



eHealth Exchange™

Q&A

Thank you for your participation.