Orientation to OCHIN data

AIM-AHEAD Program

Data & Research Core (DRC)

OCHIN

aim-ahead-drc@ochin.org

OCHIN

A driving force for health equity

Last updated: 01-22-2025

What is a Community-based Health Center (CHC)?

Health centers are community-based and patient-directed organizations that deliver comprehensive, high-quality, free or low-cost primary health care services.

(HRSA 2021. https://bphc.hrsa.gov/about-health-centers/what-health-center)

CHCs are:

- Often include pharmacy, dental, mental health, substance use disorder care
- Provide services regardless of ability to pay; charge on sliding fee scale if uninsured
- Emphasize coordinated care management, use of quality improvement practices, health information technology
- Have federal reporting requirements (UDS)

CHCs are NOT:

- Hospitals
- Health plans
- Integrated health systems

Clinics in the OCHIN network are community-based health centers. This umbrella term includes (but is not limited to):

- Federally qualified health centers (FQHCs)
- FQHC lookalikes
- Rural Health Centers
- Ryan White HIV/AIDS clinics
- Healthcare for the Homeless grantees

OCHIN Research Data Warehouse (RDW)

- OCHIN is a nonprofit leader to a growing nationwide provider network.
- OCHIN stewards the largest collection of community health EHR data in the country with more than two decades of practice-based research expertise.
- OCHIN is a founding partner of the AIM-AHEAD Data and Research Core, and the OCHIN Research Data Warehouse is the source of the AIM-AHEAD OCHIN Community Health Database.
- Data are aggregated from OCHIN's a single instance of the Epic EHR for **286** health systems with **2,463** clinic sites across **40** states

The **OCHIN RDW** integrates outpatient EHR data for patients seen in all member health centers.

- Data are standardized into a common data model based on the PCORnet CDM and will soon be available in OMOP.
- Contains all table and fields defined for the PCORnet CDM plus additional fields that are unique to safety net clinics (e.g., FPL, primary language, homeless status).

Source: OCHIN RDW, accessed 1/2025

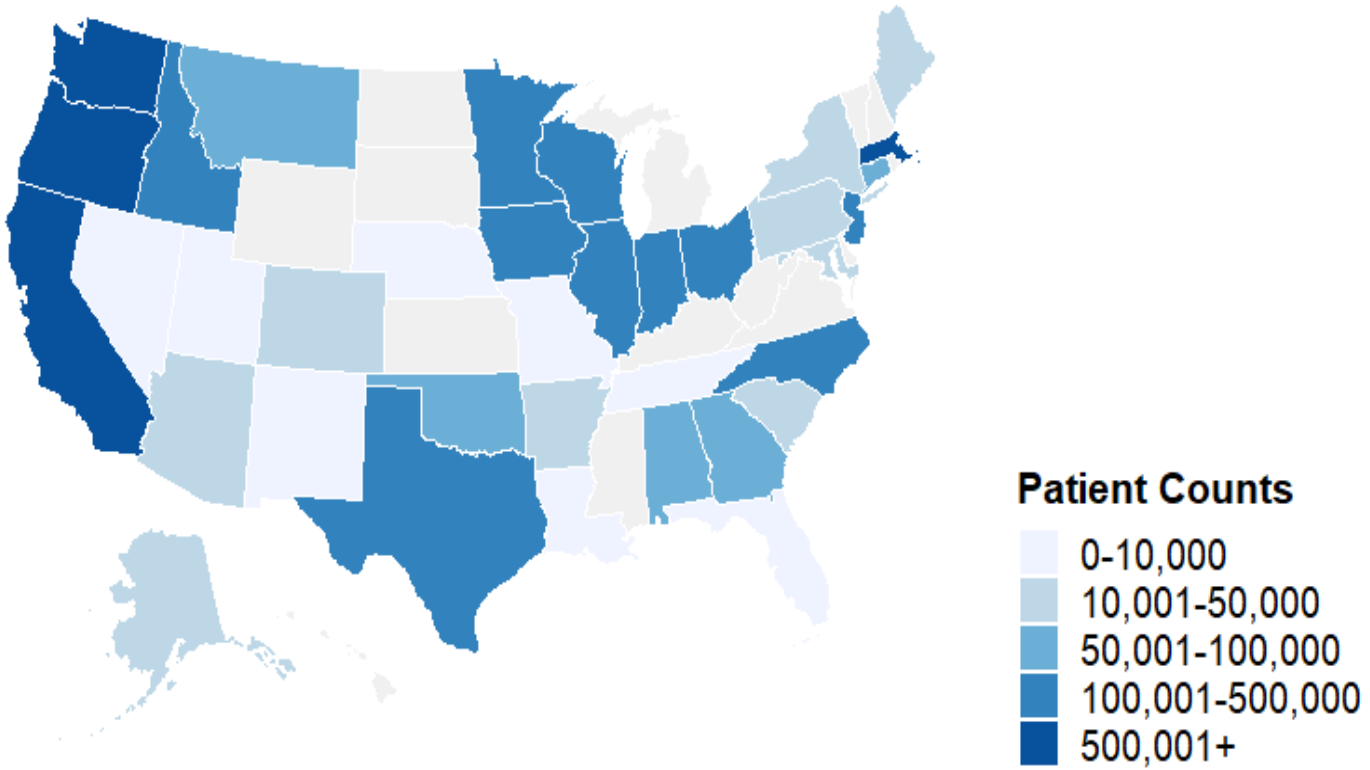
OCHIN EHR Data Overview

OCHIN operates the most comprehensive database on primary healthcare and outcomes of safety net patients in the United States.[1] The OCHIN Epic EHR data warehouse aggregates electronic health record (EHR) and social determinants of health (SDH) data representing:

- **>8** million patients (**5.5** million patients are 'active,' with a visit in the last 3 years)
- **286** health systems
- **2,463** clinic sites
- **40** states

Approved AIM-AHEAD projects can obtain access to up to 12 years of longitudinal OCHIN Epic ambulatory EHR data, which is research-ready on the PCORnet Common Data Model (CDM).

OCHIN Patients by Clinic's State



Source: OCHIN RDW, data as of 3/2024

[1] OCHIN leads and is the largest data contributor of the ADVANCE Clinical Research Network (CRN), a member of PCORnet. (https://advancecollaborative.org/)

OCHIN EHR Data

Key Characteristics of OCHIN EHR Data		
Variables	**Percent**	**Patient Count**
Total all-time patients Total active patients (seen in last three years)		8,634,916 5,591,910
100% and Below Federal Poverty Level (FPL)	40.4%	3,493,065
101% - to 200% FPL	33.4%	2,887,174
Medicare	7.5%	648,777
Medicaid	48.3%	4,171,776
Uninsured	25.7%	2,217,214
Rural	15.5%	1,342,843
Spanish Speaking	21.6%	1,861,447
Black	16.6%	1,431,134
Hispanic / Latino/a/x	33.2%	2,867,563
Asian	5.1%	444,384
American Indian/Alaska Native	1.2%	101,837
Diabetes[1]	10.7%	730,202
Hypertension[1]	20.8%	1,423,945
Asthma[1]	8.6%	585,846
Hyperlipidemia[1]	19.2%	1,309,627
Mental/Behavioral Health Dx[1,2]	38.7%	2,641,577
Obesity[3]	12.1%	824,152

[1]Chronic condition percentages presented among all-time adult patients (N=6,835,422)
[2]Includes anxiety, bipolar, depressive disorders, schizophrenia, and other psychotic disorders
[3] Obesity diagnosis on problem list or last-recorded BMI >30

Overall inclusion/exclusion of source database

- Data years available for AIM-AHEAD: 2012-2023 (>170 million total encounters)
- Patients with 1 or more ambulatory, telehealth, or dental visit at a member clinic site on or after 1/1/2012
- Records from institutionalized patients and neonates (<28 days old) are excluded.

Data Available for AIM-AHEAD Projects

Domain	Example Variables
Demographics	Sex, age, race, ethnicity, language, FPL, sexual orientation, gender identity, vital status/death date, state and zip code of residence
Encounters	Encounter type, level of service, provider type, date
Diagnoses (from encounters, problem list, and patient-reported medical history)	ICD-9 and ICD-10 diagnosis codes, description, date
Procedures (from encounters and patient-reported surgical history)	CPT and HCPCS procedure codes, description, date
Vitals	BP, BMI, and tobacco use measurements, measurement date
Laboratory results*	Lab type (standardized to LOINC), specimen source, date, result
Medications (prescribing and dispensing)*	RxNorm, NDC, medication name, dose, quantity, route, frequency, refill count
Patient-reported outcomes*	Screening questionnaire responses (e.g., PHQ2, PHQ9, AUDIT, DAST), screening date
Immunizations	Immunization type, dose, administration date
Social determinants of health*	Patient-level social needs screenings recorded in EHR, e.g., food insecurity, housing quality, housing insecurity, transportation needs, education, employment
Area-level measures	Area-level indicators at census tract and/or ZCTA level, linkable to patient addresses. Sample measures/domains: social vulnerability index, environmental justice index, GINI income inequality.

** Limitations exist for Laboratory results, Medications, Patient-reported outcomes, and Social determinants of health. See slide 18 "Highlighted Limitations" for further details.*

Unique Features of the OCHIN RDW

- Single patient record across the system, allowing unduplicated longitudinal analysis at patient level.
- Data available from 2012 through 2023.
- Many OCHIN members provide integrated primary care, dental, mental/behavioral health, and specialty care.
- Includes **>3 million** patient-level social determinants of health screens (e.g., food insecurity, housing needs).
- Includes area-level geographic data from AHRQ SDoH Database and other sources (publicly available indicators at county, census tract, and ZCTA levels).
    - Linkable to patient geocoded addresses to provide context on the environment in which patients live

Source: OCHIN RDW, accessed 1/2025

Data Not Currently Available for AIM-AHEAD Projects

Not currently available	Reason
Network member or delivery location more granular than state	Confidentiality of member clinics and patients
Chart notes	PHI disclosure risk
Family linkages within EHR*	Limited scope and completeness
PCORnet fields relating to inpatient care	Data unavailable in OCHIN CHC network
*Evaluating potential future use for AIM-AHEAD	

***NOTE**: All OCHIN data come from ambulatory clinic settings. No ED, inpatient, or sub-specialty data are available at this time.*

Notes On Using OCHIN EHR Data for Research

- Open cohort
  - All data stem from CHC being 'live' on EHR and patient utilization
  - Health centers have joined the OCHIN network at different points in time
  - Most patient data are collected within the context of an encounter
- Different systems & practices = differential data completeness and quality
- Ambulatory only – no hospitals, sub-specialties
- Vulnerable populations filtered out (e.g., incarcerated persons, neonates)
- Implications for research studies
  - Cohorts from earlier time periods will be smaller as patient counts have increased due to new facilities joining the OCHIN network. Longitudinal studies will be influenced by patient counts over time.
  - Data come from more clinics in recent years (avoid comparing raw volumes over time)
  - Watch for systematic differences by health system to avoid spurious results or conclusions

Governance

The AIM-AHEAD Community Health Database from OCHIN is a limited data set (LDS), specified in the HIPAA Privacy Rule as a dataset in which certain direct identifiers have been removed.

A limited data set is still Protected Health Information (PHI)

An LDS may include:

- Dates (e.g., date of birth, service dates)
- City, state, zip code, and/or census tract (with street address removed)

Access to data contained in this LDS requires an IRB-approved or exempt protocol and a Data Use Agreement (DUA) between OCHIN, HMS, and the requesting party.

To request access as part of an AIM-AHEAD program, start with this form.

What is a Clinic?

- CHCs have several layers of organization, which has implications in how we pull, use, and interpret data
- Different terms are used for these concepts within **conversation** and **RDW**

Health System

Organizational/ ownership level. Has a HQ address. Many health systems have multiple clinics.

Health Center, CHC
Health System

e.g., Multnomah County Health Dept

Clinic Site

Physical location, "brick & mortar". Has an address and generally contains multiple departments.

Clinic
Clinic site
Delivery site

e.g., MC East County Health Center

Generally, most relevant to research

Department

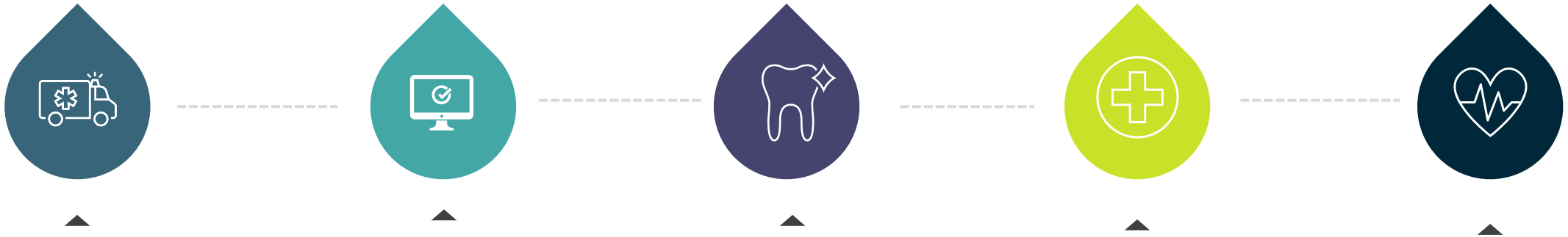
An EHR unit. Providers log into and patients have encounters at a specific departments. Many departments may be at one clinic site.

Department
Facility

e.g., Primary Care, Dental, Vaccine, Pharmacy

Notes on Key Variables

Encounter Types



Ambulatory Visit

- Majority of clinical face-to-face visits; generally billable
- Examples: primary care, well-child visits, behavioral health
- Diagnoses, procedures, labs, medications, and insurance type most often associated with AV and TH. AVs usually have vital signs (BP, BMI).

Telehealth

- Similar clinical and billable characteristics to AV, but delivered remotely
- Less likely to have vitals recorded
- Appear very rarely prior to 2020. Rapid uptake from 3/2020 onward.

Dental

Other Ambulatory

- Usually non-billable, limited-service encounters or patient contacts
- Examples: pharmacist calls, refill requests, patient registration, lab draw or immunization only
- Often filtered out of research queries

Other

- Similar to Other Ambulatory
- Often filtered out of research queries

Notes on Key Variables (Continued)

Diagnoses
(ICD-9: 2012-Sep 2015, ICD-10: Oct 2015-current)

Visit Diagnosis

- All diagnosis codes recorded in the context of encounters, usually several per encounter
- Primary dx not available

Problem List

- A list of current and active as well as past/resolved diagnoses relevant to the care of the patient. Accessible and used across healthcare team. Meant to indicate ongoing, non-transitive conditions, and/or those that are most important about a patient
- Patient-based measure, not linked to or specific to a given encounter

Medical History

- Similar concept to problem list but recorded in medical history section of chart; often patient-reported, may be more subject to recall limitations and workflow differences

Highlighted Limitations

- Not all laboratory records are mapped to LOINC
- Medication dispensing data is difficult to measure and prone to bias
  - Only exist for patients who return for a subsequent visit (captured via pharmacy data vendors and queried automatically prior to scheduled visits
  - Limited to insured patients (where Rx was paid by a public or private plan)
  - Can’t be directly linked to a prescribing record
  - Medication adherence difficult to measure (discrete days supply not available)
- Patient-reported outcomes (PRO) are not collected consistently across health systems or patient populations
- Social determinants of health (SoDH) are not collected consistently across health systems or patient populations
  - Lack of screening does not indicate absence of social need

Generalizability, Bias, and Scope of Interpretation

- Differential completeness and quality (coding differences) exist by network partner site, health system, and clinic
  - *Try to account for these differences by using the surrogate health system identifier as a clustering or control variable*
- Patient population/characteristics are not homogeneous across the OCHIN network
  - *Examine data to understand heterogeneity, identify potential sources of bias, and avoid unnecessary assumptions when making interpretations*
- Out-of-network care is captured incompletely. When care was not delivered within the OCHIN network, it could be the patient received it out of network, refused it, could not access it, or some other reason.
  - *Don't assume that care not delivered in an OCHIN clinic was not received*
- Reference/control group: still a socioeconomically disadvantaged subset of the population
  - *Please remember OCHIN's patient profile and be cautious not to generalize beyond this CHC population*