

A SHORT GUIDE TO THE ADVANCE RESEARCH DATA WAREHOUSE

LAST UPDATED: NOVEMBER 5, 2018

PURPOSE OF THIS DOCUMENT:

THIS DOCUMENT IS INTENDED TO BE A HIGH-LEVEL SUMMARY OF THE PCORNET PROJECT AND THE ADVANCE CLINICAL RESEARCH NETWORK. MORE DETAILED AND TECHNICAL INFORMATION ABOUT THE ADVANCE RESEARCH DATA WAREHOUSE AND ITS CONTENTS IS AVAILABLE IN THE “ADVANCE USER MANUAL AND DATA DICTIONARY”.

OVERVIEW OF ADVANCE:

The Accelerating Data Value Across a National Community Health Center Network (ADVANCE) Clinical Research Network (CDRN), funded by the People Centered Research Foundation (PCRF), integrates outpatient electronic health record (EHR) data for Community Health Center (CHC) patients, and integrates hospital, health plan, and community data for these patients. The CRN serves as a ‘community laboratory’ for engaging vulnerable patients in Patient Centered Outcomes Research (PCOR, <http://www.pcornet.org/clinical-data-research-networks/>).

BACKGROUND:

PCORnet was created to improve the nation’s capacity to conduct comparative effectiveness research by creating a large, highly representative electronic data infrastructure for conducting clinical outcomes research. PCORnet was originally funded by PCORI, but is now managed by the People Centered Research Foundation (PCRF) starting Nov. 1, 2018.

PCORnet consists of:

- 9 Clinical Research Networks (CRNs), such as ADVANCE
- 2 Health Plan Research Networks (HPRNs)
- 1 Coordinating Center

ADVANCE is led by OCHIN’s Practice-Based Research Network in partnership with Health Choice Network (HCN), Fenway Health, Oregon Health and Science University (OHSU), and the Robert Graham Center. We are:

- A national network of CHC systems

- Building a ‘community laboratory’ of CHCs with longitudinal outpatient EHR data on patients seen since 1/1/2012
- Forming strategic partnerships with health plans and hospitals that will bring inpatient data into existing outpatient data repository
- Bringing in community-level data to help assess environmental/external factors that impact health

All data stored in the ADVANCE Research Data Warehouse (RDW) are de-identified. Crosswalks to actual patient identifiers are kept in a secure location for patient de-duplication during data loads and for use by analytical staff as needed for approved protocols (e.g., re-identification for a clinical trial).

RESEARCH DATA WAREHOUSE (RDW) AND THE PCORNET COMMON DATA MODEL (CDM):

The ADVANCE Research Data Warehouse (RDW) is an expansion of the Common Data Model (CDM). The RDW contains all tables and fields defined for the PCORnet CDM, plus additional fields that the ADVANCE CDRN decides to include for research purposes. These fields are elements unique to CHCs that collect data for Uniform Data System (UDS) and other CHC-related reporting purposes, which are typically not collected by non-CHC clinics. These fields include Federal Poverty Level, patient primary language, migrant/seasonal worker status, homeless status, and so on. RDW data are cleaned, validated, and research-ready.

Patient inclusion criteria:

Data are included from the ADVANCE ambulatory data partners (OCHIN, HCN, and Fenway) on any patient seen in their respective systems who was defined as “active” at the time of a data extract. Patients are active if they have had a primary care, behavioral health, or dental visit in one of the networks within the past three years, as of the date of the data extract. For each active patient, all available historic data are then included in the RDW as far back in time as possible. Hospital, claims, and community vital signs data are then included for those patients.

We started extracting patients on 1/1/2015, so the RDW includes patients deemed as “active” as of 1/1/2012. Any patients that become inactive over time (i.e., were included in the RDW as “active” after 1/1/2012 but not seen in the past three years as of any given extract date) are not removed from the RDW, but are flagged as “inactive”.

DATA SOURCES:

OCHIN, Inc.: Based in Portland, Oregon, OCHIN is a health center controlled network with 110 member organizations in 20 states. These members contain:

- 610 clinics

- 1268 departments
- 9629 providers
- 5694 PC providers

Over 1 million patients are seen per year in member clinics, with over 4.6 million visits per year. In total, there are 3.3 million distinct patients in the OCHIN system database and 34.3 million visits.

- Data partners:
 - OHSU: hospital and ambulatory data on OCHIN patients seen in their system (coming in phase 2)
 - Community Vital Signs data on all OCHIN patients, supplied by the Robert Graham Center and HealthLandscape

Health Choice Network (HCN): Founded in Florida by a group of FQHCs collaborating to recover from Hurricane Andrew, HCN is similar in size to OCHIN, with a presence in 9 states.

- 505 clinics
- 24 health systems
- 1,649,381 patients

Fenway Health: Founded in 1971 as a free community clinic for the Fenway neighborhood of Boston; it now serves Boston's lesbian, gay, bisexual and transgender community, and the greater Fenway neighborhood.

- 3 clinics
- 1 health system
- 54,538 patients

UPDATING PROCESS AND FUTURE BUILD-OUT PLANS:

For each data refresh, a log table is kept to document any unexpected data changes. Data profiles of all tables and fields in the CDM are run regularly to identify potential data quality issues at the field level. We are building out the breadth of the data and we are continuing to explore expanding the RDW to advance future data needs.

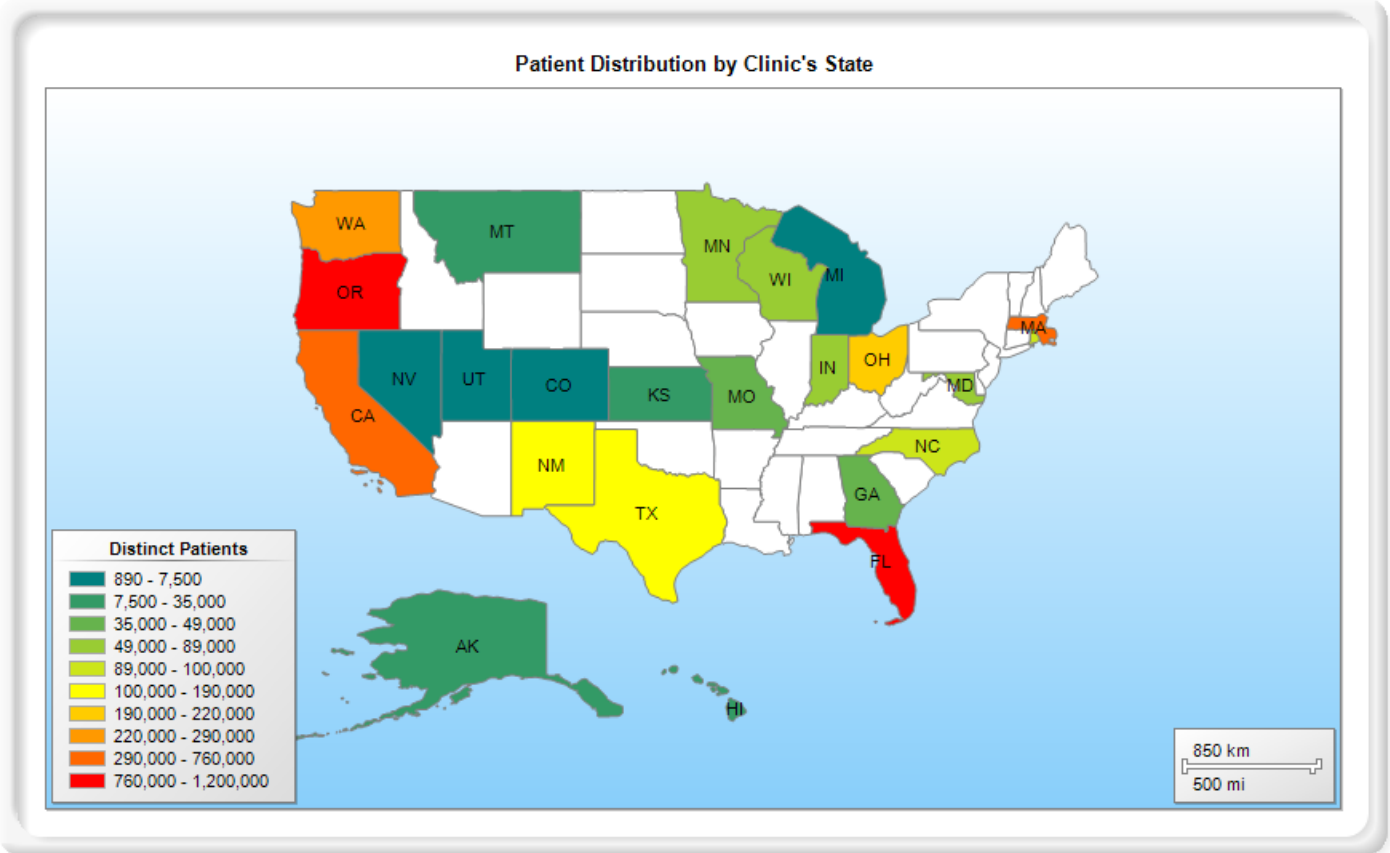
- Claims data will be added to the ADVANCE RDW in phase 2 (Sources are still to be determined)
- Community Vital Signs data on all HCN patients are planned for phase 2

OCHIN data are updated weekly, all other data (from data partners) are updated at least quarterly.

The following table lists and describes the contents of each table in the ADVANCE RDW. For a complete, detailed list of all tables and the fields within each table, see the ADVANCE RDW Data Dictionary.

Domains in RDW	Description	Fields Specific to Domains
Community Vital Signs	Aggregate data of social and environmental factors that may impact patient health.	College graduates, Median income, Professional jobs, FPL levels, Unemployment, Race, Residential segregation, Age dependency ratio, Water access, Fast food availability, Liquor store availability, Population density, etc.
Condition	Representation of a patient's diagnosed and self-reported health conditions and diseases. Includes medical history and current state.	Condition ID, Report and Resolve date, Onset date, Status, Condition code, Type, and Source, etc.
Death	Reported mortality information for patients.	Source, etc.
Death Cause	The individual causes associated with a reported death.	Death cause, Code, Type, Source, and Confidence
Demographic	Current (most recently available patient information at time of data refresh) direct attributes of individual patients with at least one medical ambulatory encounter on or after 1/1/2012. No infants younger than 31 days at time of data extract.	Birth date and time, Sex, Race, Language, Family size, Income, FPL, Payor, Migrant and Homeless Status, Veteran Status, Sexual Orientation, Gender Identity, Charlson score, etc.
Diagnosis	Indication of the results of diagnostic processes and medical coding within healthcare delivery.	Diagnosis ID, Encounter type, Diagnosis code, type, and source
Dispensing	Outpatient pharmacy dispensing, such as prescriptions filled through a neighborhood pharmacy with a claim paid by an insurer.	Dispensing ID, Prescribing ID, NDC, Supply, Amount, etc.
Encounter	Description of the interactions between patients and providers within the context of healthcare delivery.	Encounter type, Admit and Discharge dates, Provider type, Facility location and ID, Discharge status, DRG, Migrant and Homeless status, Admitting source, Income, FPL, Payor, etc.
Enrollment	Designed to identify periods during which a person is expected to have complete data capture for one enrollment period.	Enrollment start date, end date, etc.
Family Linkage	Contains mother's patient ID where a relationship has been determined by one or more data sources and algorithms.	Mother's guarantor, Coverage, Medicaid case number, Address, etc.
Immunization	Immunizations ordered and administered within healthcare delivery. Contains historical, patient-reported immunizations administered elsewhere.	Immunization date, CVX code, CPT code, Route, Order date, Vaccine lot number, Expiration date, etc
Lab Result	Quantitative and qualitative measurements from blood and other body specimens. These standardized measures are defined the same way across all PCORnet networks.	Lab result ID, Lab name, Specimen source, LOINC, Priority, Result, Procedure code and Type, Order and Specimen date, Result date and Time, Qualitative and Quantitative result, Unit, Range, etc.
Patient Reported Outcome Common Measures	Standardized measures that are defined in the same way across all PCORnet networks. Recorded at individual item level: question/statement, paired with its standardized response options.	PRO CM ID, PRO item, PRO LOINC, Response, Method, Mode, CAT, Code, etc.
Prescribing	Provider orders for medication dispensing and/or administration.	Prescribing ID, Order date and Time, Quantity, Refills, Days supply, Frequency, Basis, etc.
Procedures	Description of the discreet medical interventions and diagnostic testing, such as surgical procedures, administered within healthcare delivery.	Procedure ID, Encounter type, Procedure code and type, etc.
Surgical History	Patient's self-reported history of past surgical procedures.	Surgical history ID, History date, Procedure date, Code, Type, and Source
Vital	Vital signs (height, weight, and blood pressure) to represent a patient's current state.	Vital ID, Source, Height, Weight, BP, BMI, Smoking, Tobacco use and Type, etc.

Appendix A: ADVANCE Patient Distribution by Clinic's State



Health Systems	Clinic Sites	Cities	States
134	1116	339	24

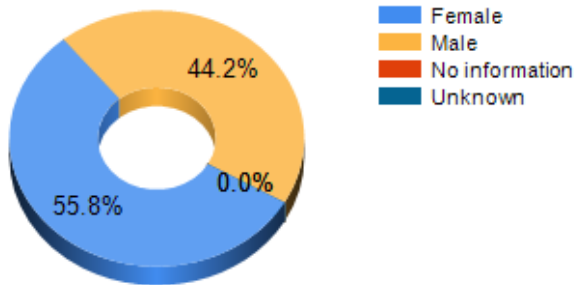
Appendix B: ADVANCE Selected Patient Characteristics

Patients by sex and selected condition:

Sex	Patients	%
Female	2,526,511	55.77%
Male	2,003,119	44.21%
Unknown	525	0.01%
No information	408	0.01%
Total	4,530,563	100.00%

Patients 18 and older by Condition		
Condition	Patients	%
Diabetic (Type I or II)	313,875	9.08%
Diabetic Type I	13,600	0.39%
Diabetic Type II	300,275	8.68%
Obese	1,084,462	31.36%
Obese & Diabetic (Type I or II)	177,636	5.14%
Alpha-1	271	0.01%
Total Adults	3,458,275	100.00%

Patients by Sex

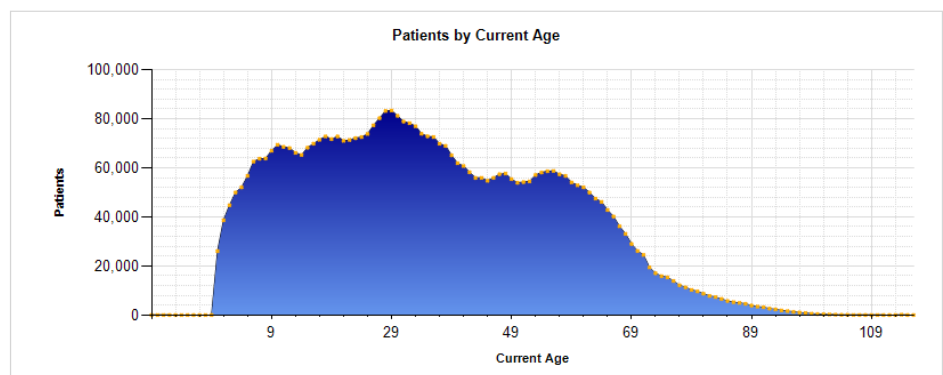


Patients by race and ethnicity:

Race \ Hispanic Ethnicity	No	Yes	Unknown	No information	Refuse to answer	Total	%
White	1,539,691	1,233,870	35,925	55,233	31,642	2,896,361	63.9%
Black or African American	797,629	44,265	11,529	19,635	23,009	896,067	19.8%
Asian	131,707	3,489	3,746	6,498	4,242	149,682	3.3%
Unknown	43,921	102,121	94,232	190	768	241,232	5.3%
Refuse to answer	28,814	87,802	919	2,172	12,200	131,907	2.9%
Multiple race	34,921	28,541	801	2,388	2,513	69,164	1.5%
No Information	15,663	33,840	4,276	28,491	4,348	86,618	1.9%
Native Hawaiian or Other Pacific Islander	19,046	5,767	577	288	1,877	27,555	0.6%
American Indian or Alaskan Native	16,986	11,680	594	561	495	30,316	0.7%
Other	120	95	38	1,408		1,661	0.0%
Total	2,628,498	1,551,470	152,637	116,864	81,094	4,530,563	100.0%
%	58.0%	34.2%	3.4%	2.6%	1.8%	100.0%	

Patients by current age:

Age Category	Patients	%
Unknown	11	0.0%
-30--21	10 or less	0.0%
-20--11	10 or less	0.0%
-10--1	10 or less	0.0%
0-9	525,275	11.6%
10-19	691,535	15.3%
20-29	757,856	16.7%
30-39	738,508	16.3%
40-49	573,851	12.7%
50-59	563,336	12.4%
60-69	429,751	9.5%
70-79	165,831	3.7%
80-89	64,354	1.4%
90-99	18,739	0.4%
99+	1,497	0.0%
Total	4,530,563	100.0%



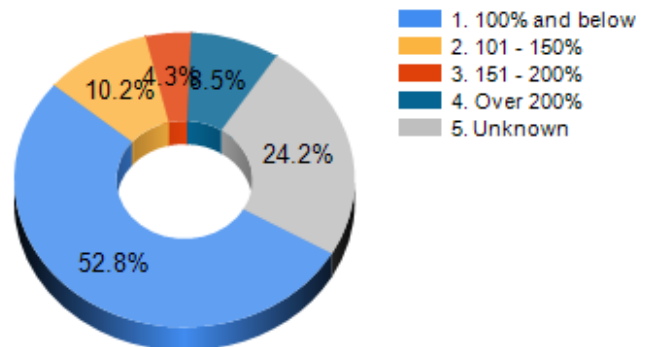
Patients by primary language:

Patients by primary language - Top 10		
Primary Language	Patients	%
English	3,307,506	73.0%
Spanish	889,247	19.6%
Unknown	154,917	3.4%
Other language	86,059	1.9%
Somali	23,605	0.5%
Vietnamese	18,903	0.4%
Arabic	14,285	0.3%
Russian	13,690	0.3%
Creole	9,172	0.2%
Nepali	7,538	0.2%
French Creole Haitian	5,641	0.1%
Total Patients	4,530,563	100.0%

Patients by their most recent federal poverty level:

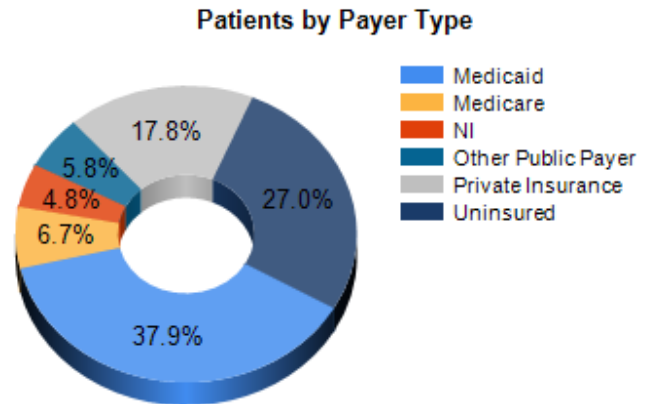
FPL Category	Patients	%
1. 100% and below	2,390,744	52.8%
2. 101 - 150%	462,773	10.2%
3. 151 - 200%	195,227	4.3%
4. Over 200%	383,447	8.5%
5. Unknown	1,098,372	24.2%
Total Patients	4,530,563	100.0%

Patients by FPL



Patients by payer type on most recent visit:

Payer Type	Patients	%
Medicaid	1,718,571	37.9%
Uninsured	1,222,141	27.0%
Private Insurance	807,953	17.8%
Medicare	303,535	6.7%
Other Public Payer	263,142	5.8%
NI	215,221	4.8%
Total Patients	4,530,563	100.0%



For further reports and details on the patient population, please refer to the link below:

http://research_sql/Research/Pages/Report.aspx?ItemPath=%2fADVANCE+Data+Profiles%2fADVANCE+Demographics+by+Site