

Real World Data for Research in the GPC

PCORnet CDM and the
Greater Plains Collaborative
(GPC)

Welcome to the Greater Plains Collaborative (GPC)

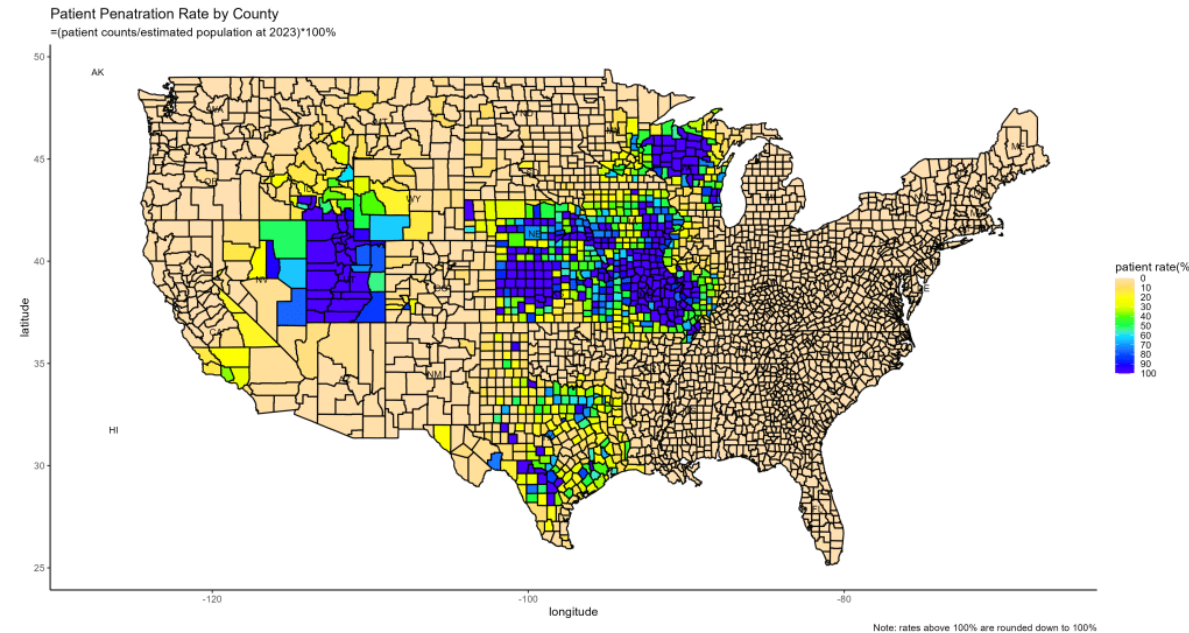
The **GPC** is one of the 8 PCORnet's regional networks

The GPC includes 14 major health systems

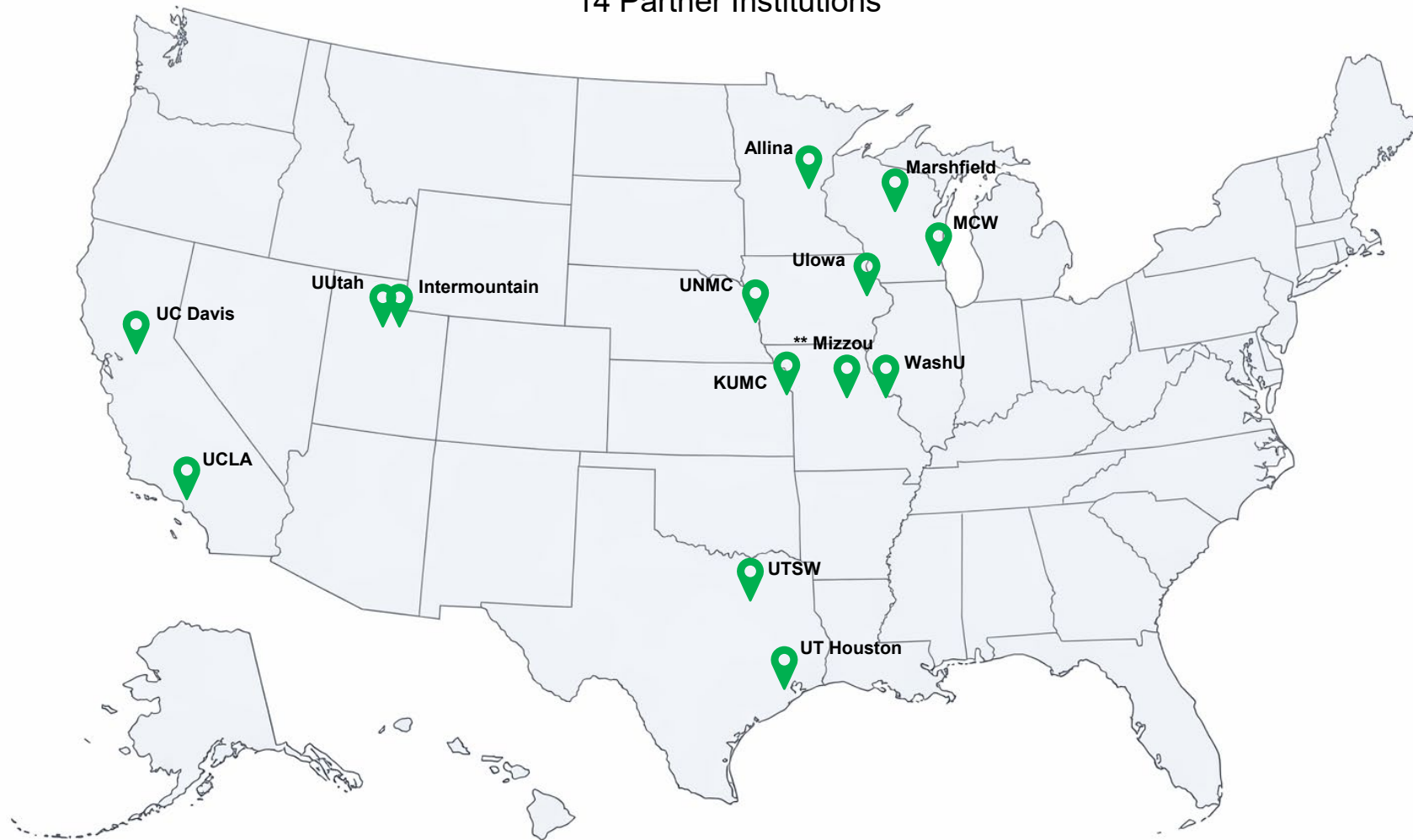
GPC Strengths:

- Deep EHR data
- Longitudinal patient histories
- Strong governance
- High data quality
- Expertise in distributed research

GPC sites transform their local EHR data into the **PCORnet CDM**, enabling cross-site studies.



14 Partner Institutions



What is Real World Data (RWD)

- Data generated during routine clinical care
- Not collected in controlled trials
- Examples:
 - Electronic health records (EHRs)
 - Insurance Claims and billing data
 - Lab results, vitals, medications
 - Diagnoses and procedures
 - Patient-reported outcomes

RWD is powerful because it reflects actual clinical practice

The Evolution of the Real-World Data in Research

The Age of Restriction

Bench to Bedside

Meaningful Use

Breaking down barriers

The future is here...

1996 - 2005



Health Insurance
Portability and
Accountability Act of 1996

Protected Health
Information
HIPAA Privacy Rule 2000

HIPAA Security Rule 2003
ePHI protections

2005-2010



Clinical and
Translational Science
Award (CTSA)
program



NCATS

Established 2011

2011-2020



21st Century
Cures Act (2016)

ONC Cures Act
Final Rule 2020

THE LEARNING
HEALTHCARE PROJECT

2020 - 2026

EHR access through
APIs “without special
effort”



2026-

AI Assisted data
recording and
analysis

Personal Devices

Where we ask,
“Why can’t people
collect their own
data?”



Why RWD Matters

- Enables large, diverse patient cohorts
- Supports longitudinal follow-up
- Captures real-world clinical variation
- Faster and more cost-effective than traditional trials
- Essential for comparative effectiveness and patient-centered research

What Is PCORnet data?

PCORnet (the National Patient-Centered Clinical Research Network) is a national research infrastructure funded by PCORI.

Its purpose is to make large-scale, multi-site research **faster, cheaper, and more reliable** by using RWD.

PCORnet includes:

- 70+ health systems
- Complete records on over 40 Million patients
- Shared data standard: **PCORnet Common Data Model (CDM)**
- Distributed research model (data stays local)

Face-to-Face Encounter Care Setting

Ambulatory Visit	40,539,250 (87%)
Emergency Department	9,213,321 (20%)
Telehealth	6,360,676 (14%)
Inpatient Hospital Stay	2,991,480 (6%)
Emergency to Inpatient	1,034,540 (2%)
Observation Stay	597,139 (1%)
Non-Acute Institutional Stay	89,990 (<1%)

What Is the PCORnet Common Data Model (CDM)?

The **PCORnet CDM** is a standardized way of organizing clinical data so that every participating site uses the **same table names, same variables, same formats**.

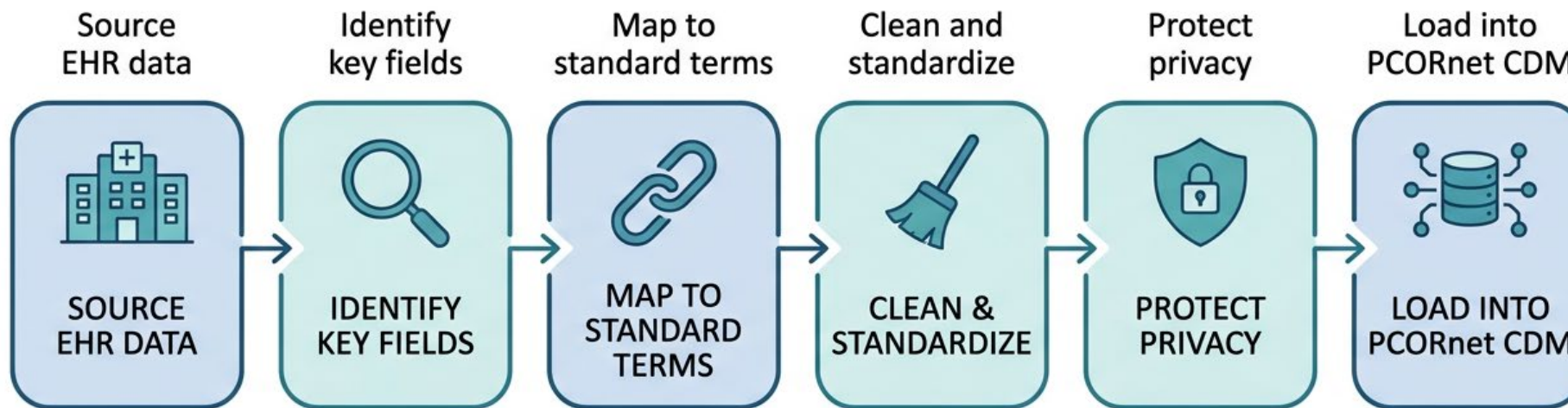
Why this matters:

- A researcher can ask a question once
- Every site can run the same query
- Results can be combined without messy data cleaning

Example:

- One site may store birthdate as DOB, another as Birth_DT.
- The CDM standardizes it to BIRTH_DATE everywhere.
- This makes multi-site research scalable and reproducible.

Converting Data from a Source EHR into PCORnet Common Data Model for Clinical Research



Data Source from

- Electronic Health Records
- Claims data
- Disease Registries
- Ancillary Systems

PCORnet Common Data Model

- Seventh version
- Over 200-page specification
- Details table design, allowed values, standard terminology

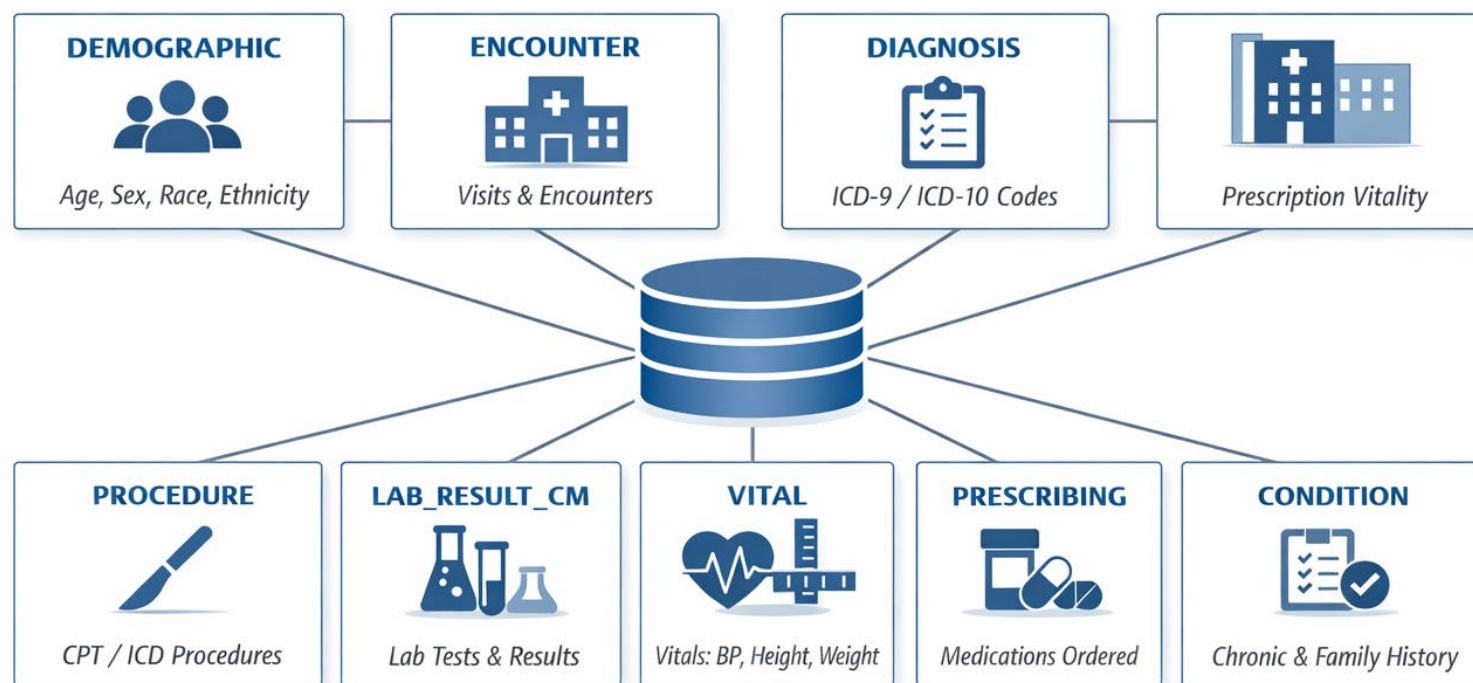
Quality
Check



Empirical
Data
Characteriza
tion Reports

Query
through
PCORnet
Front Door

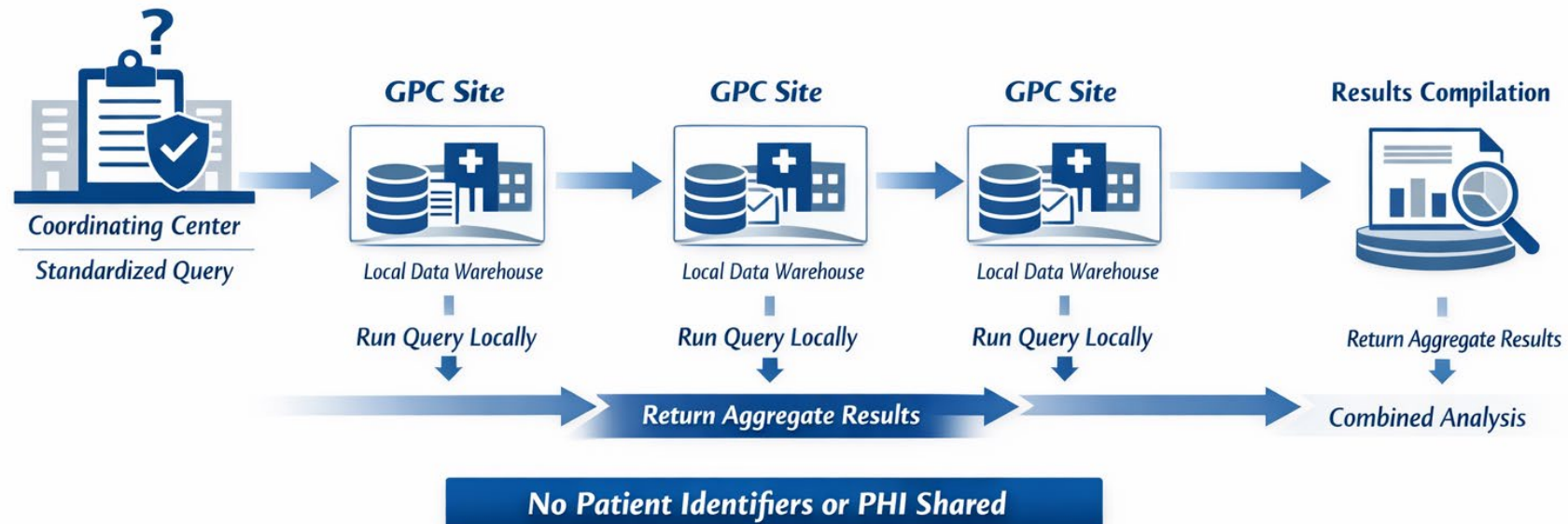
PCORnet CDM



PCORnet's Distributed Query/Research Model

- Sites keep their data locally
- A single query is written centrally
- Each site runs the same query
- Sites return aggregate results (counts, stats)
- No PHI leaves the institution

Distributed Query Model



Key PCORnet CDM Tables

- DEMOGRAPHIC — age, sex, race, ethnicity
- ENCOUNTER — visits, inpatient/outpatient
- DIAGNOSIS — ICD-9/10 codes
- PROCEDURE — CPT/ICD procedures
- LAB_RESULT_CM — labs, units, values
- VITAL — height, weight, BP
- PRESCRIBING — medication orders
- CONDITION — chronic conditions, problem lists

Example: Mapping a Research Question to CDM Tables

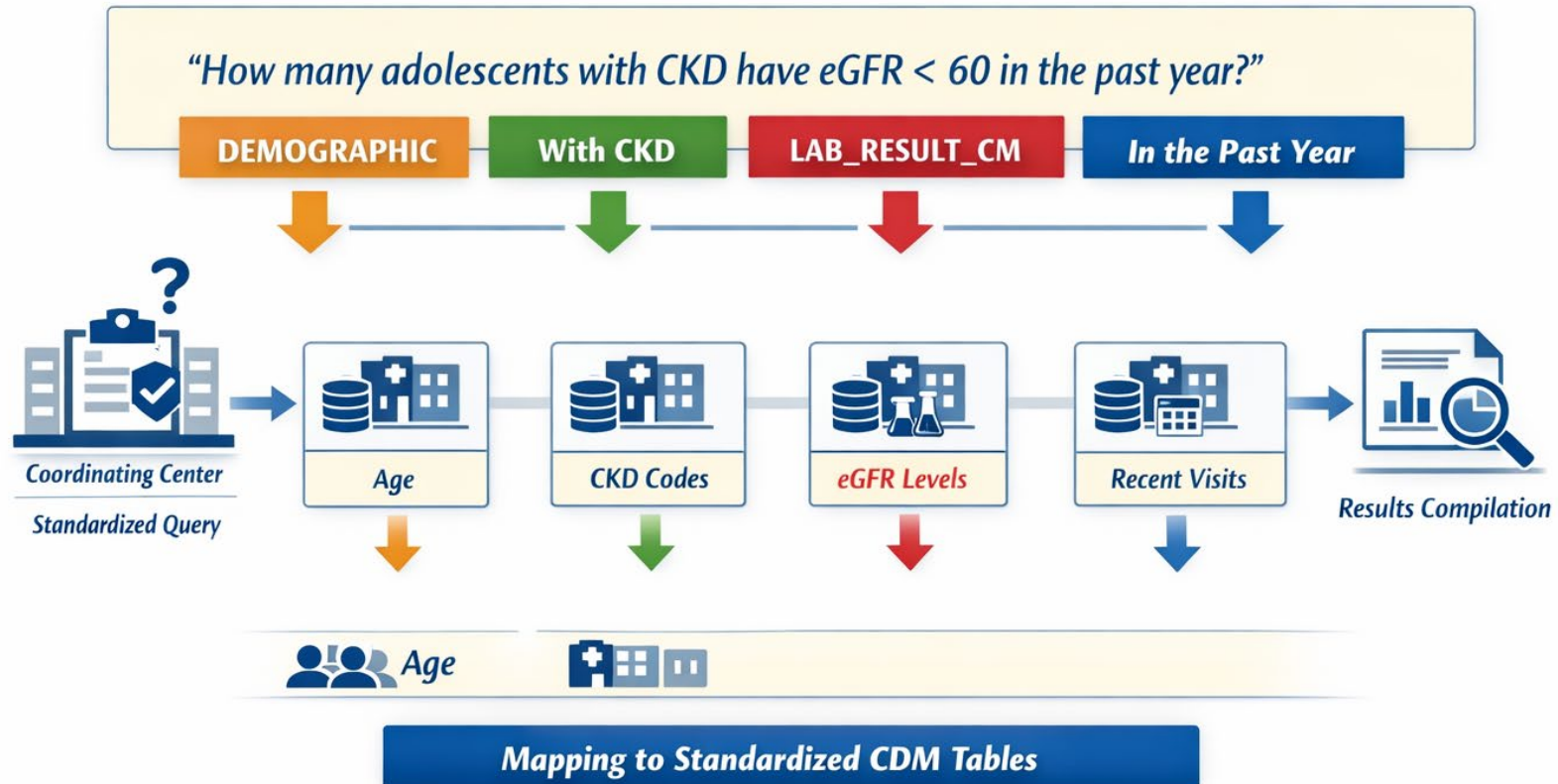
Question:

- “How many adolescents with CKD have eGFR < 60 in the past year?”

CDM Mapping:

- DEMOGRAPHIC → age
- DIAGNOSIS → CKD codes
- LAB_RESULT_CM → creatinine, eGFR
- ENCOUNTER → visit dates

Mapping a Research Question to CDM Tables



Real-World Data (RWD) Research Workflow



GROUSE: GPC Research Portal

Purpose: Unified search environment for GPC research data

Users: Investigators, analysts, and coordinators across GPC sites

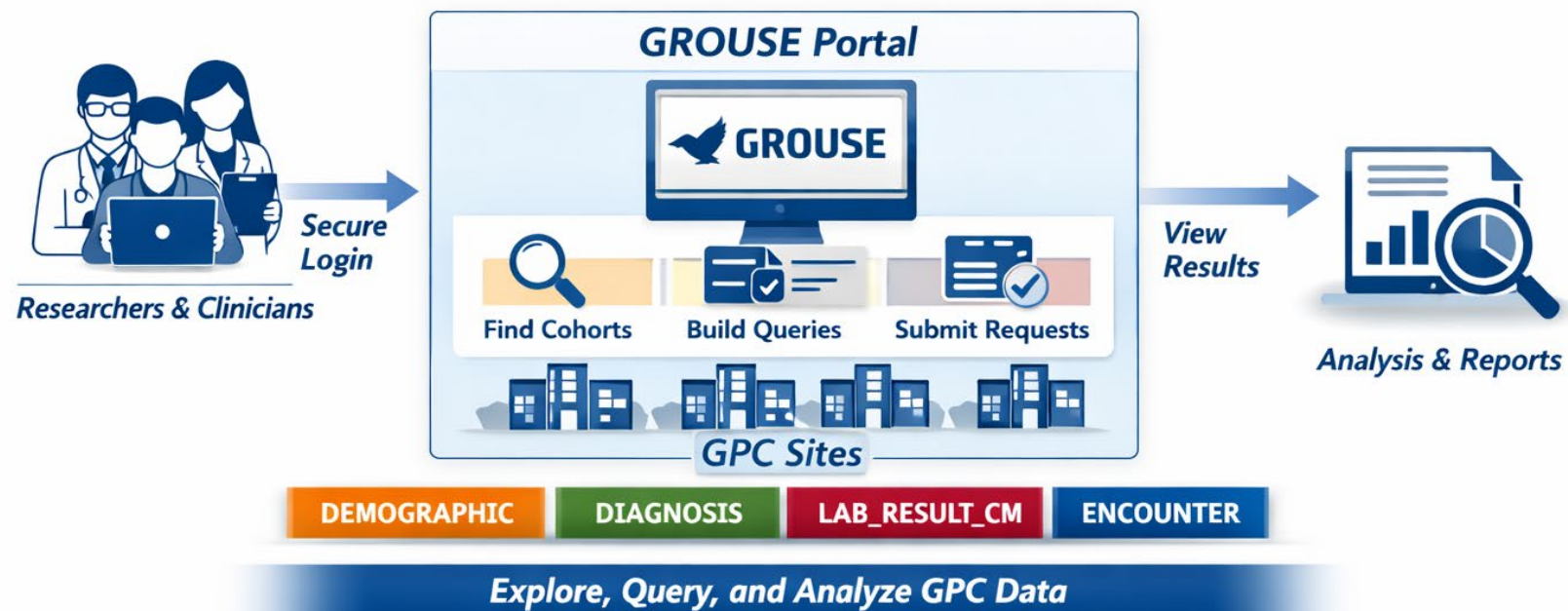
Functions:

- Explore available datasets and cohorts
- View site-level data availability
- Submit data requests and queries
- Track study progress and data refresh cycles

Governance: Access controlled via GPC Data Governance Committee

Integration: Built on PCORnet CDM structure; supports distributed queries

GROUSE: GPC Research Portal



Benefits for New Researchers

- No need to handle Protected Health Information (PHI)
- Standardized data reduces complexity
- Large sample sizes
- Faster study startup
- Strong governance and support
- Ideal environment for learning RWD research

Example Use Cases

- Chronic disease surveillance
- Medication safety
- Health disparities research
- Pediatric and adolescent health
- Genomics + clinical data integration
- Comparative effectiveness studies

PCORnet Data Insights

Prespecified Disease Population Queries

Provide a sense of scale for research projects



<https://pcornet.org/data/population-insights/>

Summary

- RWD enables real-world, patient-centered research
- PCORnet CDM standardizes data across the country
- GPC provides high-quality, multi-site research infrastructure
- New staff can quickly become productive using CDM-based workflows

Questions

- Can we put data collection tools in hands of participants?
- Should EHR vendors support research use case?