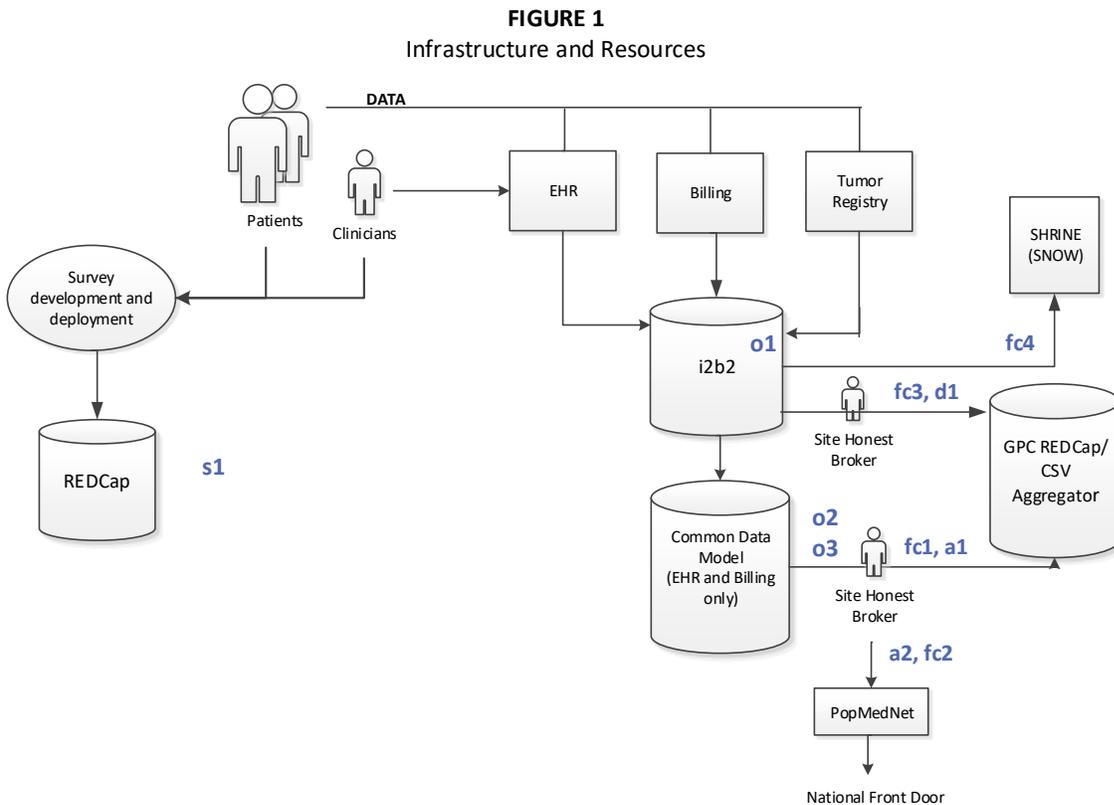




GPC Resource Guide for Investigators

The GPC brings together the biomedical informatics infrastructures and resources from twelve leading medical centers and teaching hospitals. Additionally, as a Clinical Data Research Network (CDRN) within PCORnet, the GPC can work collaboratively on data requests and research initiatives with similar networks nationally. This work is facilitated through the PCORnet Coordinating Center which operates PCORnet’s “Front Door”.

As you plan your next research study, explore the GPC’s multiple opportunities for potential collaboration. The following figures and supporting descriptions provide an overview of our resources and services.



Tools Available for Investigator Orientation

- Babel - Babel provides investigators with a list of available data from all GPC sites’ i2b2 instances (Figure 1 - o1). Request a log-in and explore at babel.gpcnetwork.org.
- Electronic Data Checks (EDC) – GPC sites execute standardized data characterization reports developed nationally by the PCORnet Coordinating Center. These reports provide insights on the amount of data and its quality from each GPC site (Figure 1 - o2) - *Available Late Spring 2017*

- Demographics – Total enrollment plus other key demographic data by site (Figure 1 - o3). Available upon request.

Feasibility Counts

- Common Data Model (CDM) - PCORnet created the CDM for use among its participating networks to standardize electronic health record and billing data in support of observational studies and clinical trials. Data from the CDM can be made available at either the GPC site level (Figure 1 - fc1) or from multiple Clinical Data Research Networks (pcor.net/participating-networks), obtainable through inquiries to the PCORnet Front Door (Figure 1 -fc2): pcor.net/frontdoor.

The CDM has evolved through three versions to its current version of v3.1. The CDM domains are summarized below.

PCORnet Common Data Model Domains, v3.0 and v3.1

DEMOGRAPHIC v1.0

Demographics record the direct attributes of individual patients.

ENROLLMENT v1.0

Enrollment is a concept that defines a period of time during which a person is expected to have complete data capture. This concept is often insurance-based, but other methods of defining enrollment are possible.

ENCOUNTER v1.0

Encounters are interactions between patients and providers within the context of healthcare delivery.

DIAGNOSIS v1.0

Diagnosis codes indicate the results of diagnostic processes and medical coding within healthcare delivery. Data in this table are expected to be from healthcare-mediated processes and reimbursement drivers.

PROCEDURES v1.0

Procedure codes indicate the discrete medical interventions and diagnostic testing, such as surgical procedures and lab orders, delivered within a healthcare context.

VITAL v1.0

Vital signs (such as height, weight, and blood pressure) directly measure an individual's current state of attributes.

LAB_RESULT_CM v2.0

Laboratory result Common Measures (CM) use specific types of quantitative and qualitative measurements from blood and other body specimens. The common measures are defined in the same way across all PCORnet networks, but this table can also include other types of lab results.

CONDITION v2.0

A condition represents a patient's diagnosed and self-reported health conditions and diseases. The patient's medical history and current state may both be represented.

PRO_CM v2.0

Patient-Reported Outcome (PRO) Common Measures (CM) are standardized measures that are defined in the same way across all PCORnet networks. Each measure is recorded at the individual item level: an individual question/statement, paired with its standardized response options.

DISPENSING v2.0

Outpatient pharmacy dispensing, such as prescriptions filled through a neighborhood pharmacy with a claim paid by an insurer. Outpatient dispensing may not be directly captured within healthcare systems.

PRESCRIBING v3.0

Provider orders for medication dispensing and/or administration. These orders may take place in any setting, including the inpatient or outpatient basis.

PCORNET_TRIAL v3.0

Patients who are enrolled in PCORnet clinical trials.

DEATH v3.0

Reported mortality information for patients.

DEATH_CAUSE v3.0

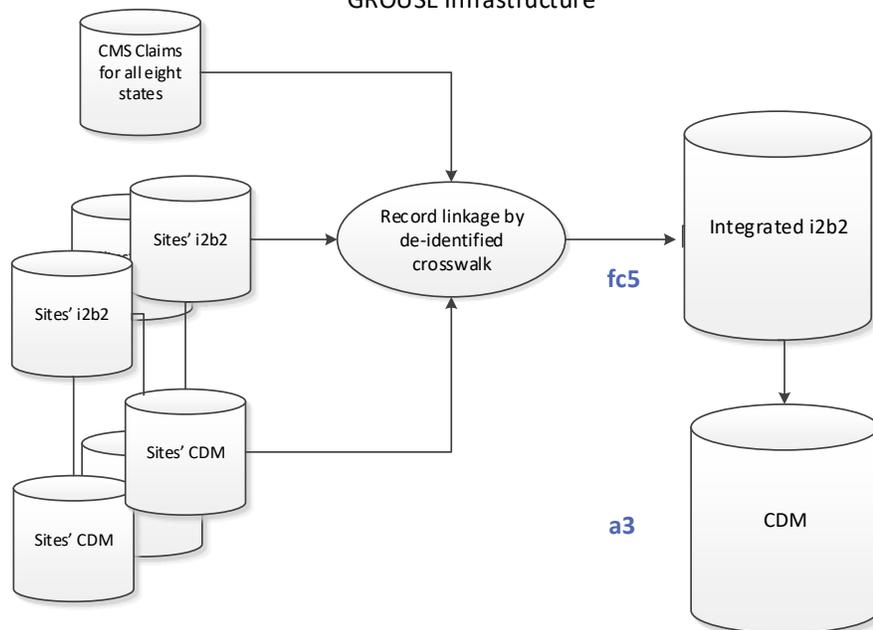
The individual causes associated with a reported death.

HARVEST v3.0

Attributes associated with the specific PCORnet datamart implementation, including data refreshes.

- i2b2 + Cancer Registries (Figure 1 – fc3) - i2b2 (i2b2.org), a software framework for integrated data repositories widely used across the NIH Clinical and Translational Science Award institutions, is deployed and updated at all GPC sites. It allows self-service querying and also data extraction in a consistent manner. Notably, all GPC sites have incorporated their standardized hospital tumor registries alongside their billing and electronic health record data in their i2b2s which are available for your analysis and retrieval.
- SHRINE (Shared Health Research Informatics Network) is a software system that allows GPC investigators to automatically query across sites' i2b2 instances to obtain feasibility queries without intervention by the honest brokers at each site. For GPC, our leaders in Wisconsin manage this network as the Shrine Network of Wisconsin (SNOW) for self-service feasibility counts across all GPC i2b2 instances (Figure 1 - fc4)- Available late Spring 2017 .
- GROUSE (Greater Plains Collaborative Reusable Observable Unified Study Environment) is a project to merge the Medicare/Medicaid data for the GPC's eight states with the data from each of our sites EHR and billing data provided via i2b2 or their CDM tables. It can be used currently to analyze our three initial cohorts: ALS, Breast Cancer or the impact of Height/Weight on health. It may also be repurposed for other conditions with CMS approval (Figure 2 - fc5). This environment currently has 3 years of claims (with plans to purchase 2 more years) for the entire states with 3.8 billion facts covering 19 million beneficiaries including hospitalization, outpatient professional procedures and diagnoses, home health and hospice services, and Medicare Part D drug benefits and claims organized by National Drug Codes.

Figure 2
GROUSE Infrastructure



Data

- Investigators may request either de-identified or limited data sets* from i2b2 repositories which will be integrated into their project specific REDCap database (secure web application for building and managing online surveys and databases). The raw data is provided via csv files (Figure 1 - d1)

Analysis

- Investigators can provide SAS code to execute against site(s) CDM returning aggregate results.
 - Distributed thru GPC (Figure 1 - a1)
 - Distributed thru PCORnet Front Door (Figure 1 - a2)
- Investigators can provide SAS or R code to execute in the GROUSE environment against i2b2, CMS Claims, and CDM formats (Figure 2 - a3)
 - If your work is aligned with the current GPC cohorts of ALS, Breast Cancer or Height/Weight → results can be returned without additional approval from CMS
 - Non GPC cohort topics → Results from GROUSE* (will require data reuse approval from CMS and GPC staff will assist with this in coordination with RESDAC at the University of Minnesota)

Surveys

- Investigators can develop surveys via REDCap to provide preliminary data regarding Patient Clinician Engagement for their studies. After they develop their REDCap Data Dictionaries, GPC sites can deploy these surveys supported by Informatics' teams and Patient Engagement Officers. The survey responses are aggregated and returned to Investigators (Figure 1 - s1)

*IRB and Data Request Oversight Committee approvals required for limited data sets.

Note: To streamline limited datasets and survey deployment all GPC sites participate in the national SMART IRB initiative (smartirb.org).

For more information, visit gpcnetwork

Email: gpcnetwork@kumc.edu