Data Domains in PEDSnet database

Domain	Examples	Utility to research
Demographics	Age, date of birth, gestational age, sex, ethnicity, race, zip code, PEDSnet health system site, primary care provider (PCP), death and cause of death (if available), tobacco use	Exposures, health determinants, confounders, mediators, effect modifiers, outcomes (death); Zip codes and health system site can be geocoded and linked to area-level environmental factors influencing health and health outcomes;
Outpatient encounters (~75 visit specialty types available)	Primary care visits, specialty care clinics (e.g., cardiology, endocrinology, nephrology, oncology, etc), physical therapy, occupational therapy, speech language pathology, medical genetics/genomics, etc	Provides longitudinal follow-up and duration of clinical course; health care utilization measures; Distinguishes between specialty care (and type of specialty care) and primary care visits;
Inpatient admissions	Length of stay, discharge status; diagnoses, procedures, medications, and lab results associated with inpatient stay	Measures health care utilization for patients and conditions; Provides longitudinal data to evaluate disease course and severity of illness;
Emergency department encounters	Diagnoses, procedures, medications, and lab results associated with ED visit; ED visits resulting in inpatient admission	Provides markers for disease progression and clinical outcomes; Health care utilization measures; Markers for chronic disease control and severity;
Anthropometrics	Height (cm) and weight (kg), BMI, head circumference	Outcomes, health determinants, confounders, mediators, effect modifiers; Markers for health status;
Vital signs	Temperature, blood pressure	Outcomes, health determinants, confounders, mediators, effect modifiers; Markers for health status
Providers	Specialty, health care facility	Differences and similarities in clinical course of care between patients seen by primary care and specialty care physicians;
Diagnoses	Final diagnoses are recorded at each encounter and are mapped to a standardized codes (typically SNOMED-CT)	Provides ability to define and construct cohorts and outcomes of interest; Sufficient power to study rare diseases of interest; Provides comorbidity information;
Procedures	Procedures are recorded at each encounter and are mapped to standardized SNOMED-CT, ICD-9 Procedure, ICD-10 Procedure, CPT-4, and HCPCS codes	Measures disease severity; Improves case-finding accuracy and provides more detailed clinical data;
Prescribed medications	Medications ordered are recorded at each encounter and mapped to a standardized RxNorm code	Provides information about disease control and severity; Markers of clinical course and health behaviors;
Dispensed medications	Medications dispensed are recorded and mapped to a standardized RxNorm code	Medications that patients likely received; More accurate measure of medication record and disease control
Laboratory test results	Key components of lipid panel, complete metabolic panel, complete blood counts, microbiologic cultures, liver function tests, urinalysis, viral panels, etc	Identifies more precise clinical outcomes and health determinants; Evaluate and define disease progression; Provides more accurate measure and definition of disease as well as disease severity and control;
Visit Payer	Plan class (private/commercial, medicaid/sCHIP, Medicare, other public, self-pay, other/unknown), plan type (HMO, PPO, POS, fee for service, other/unknown) **for every unique visit.** Enrollment information unavailable.	Confounders, mediators, effect modifiers; Ability to study effect of payer visit with health outcomes and other outcomes of interest

The **Demographics Domain** contains demographic data at the person-level for all patients in the PEDSnet database. Example variables include age, date of birth, gestational age, sex, ethnicity, race, zip code, PEDSnet health system site, primary care provider (PCP), death and cause of death (if available), and tobacco use. Demographic data is a key element in all clinical health research and used to measure exposures, health determinants, confounders, mediators, effect modifiers, and health outcomes (e.g., death or cause of death). Researchers can also geocode zip codes and health system sites and study the link between environmental factors and health outcomes.

The **Outpatient Encounters Domain** captures all outpatient encounters for patients in our database who attend a site-affiliated office or institution. The PEDSnet database identifies ~75 different visit specialty types. Examples of specialty visits include primary care offices, specialty clinics (e.g., cardiology, endocrinology, nephrology, oncology, etc), development and behavioral appointments (e.g., physical therapy, occupational therapy, speech language pathology), and other miscellaneous visits (e.g., medical genetics/genomics). The Outpatient Encounters Domain provides researchers with longitudinal follow-up and duration of a patient's clinical course of care as well as health care utilization measures. Information in this domain includes dates of service, diagnoses, procedures, and information about physician and location specialty.

Similar to the Outpatient Encounters Domain, the **Inpatient Encounters Domain** encompasses all inpatient encounters for patients in our database who attend a PEDSnet site-affiliated hospital. Examples of data included in this domain are length of stay and discharge status, and all diagnoses, procedures, medications, and lab results linked to a unique inpatient visit. This domain is of high utility to researchers. Inpatient medical encounters can serve as measures for healthcare utilization and studied by disease type (e.g., oncology patients) or patient profile (e.g., patients in geographic locations or age groups). The Inpatient Encounters Domain also provides researchers with longitudinal data to evaluate disease course and severity of illness. The application of this domain can extend to almost any type of epidemiologic or clinical health study.

Analogous to the Outpatient Encounters Domain and the Inpatient Encounters Domain, the **Emergency Department Encounters Domain** consists of all diagnoses, procedures, medications, and labs associated with an ED visit. Similar measures of health care utilization and clinical course of care can be applied to this domain. Additionally, we can readily identify emergency visit encounters that result in inpatient admissions in the PEDSnet database.

The **Anthropometrics Domain** contains patient-level information regarding height (cm) and weight (kg) and head circumference. This domain serves a similar function as the **Vital Signs Domain**, which encompasses all temperature and blood pressure measurements for patients in the PEDSnet database. Together, these domains provide researchers with data for outcomes and health determinants in a study, as well as for confounders, mediators, and effect modifiers in statistical analyses. The variables in these two domains serve as markers of general health status.

The **Providers Domains** encompasses person-level information for any health care professional who provides care for patients enrolled in the PEDSnet database. This domain contains information regarding provider specialty as well as associated health care facility. The data can serve a variety of functions in health care research, such as examining the differences and similarities in the clinical course of care between patients seen primarily by their primary care doctors versus those seen by specialty care physicians. It also provides the ability to flag visits and profile patient characteristics based on specialty care. There are currently > 120 unique physician specialty types encompassed in the database.

The **Diagnoses Domain** encompasses all diagnostic information for patients in the database (>32,000 unique conditions captured). Final diagnoses are recorded at each encounter and mapped to a standardized code, typically SNOMED-CT. Because diagnostic codes are captured at every encounter, this domain provides researchers with the unique ability to longitudinally assess the duration of a patient's illness. Almost all health care research hinges on the ability to accurately assess patient diagnoses and duration of illness. The diagnoses domain provides researchers with the ability to define and construct cohorts and outcomes of interest as well as accurately assess comorbidities. Because of its size and the high number of patients enrolled in our database, studies will be sufficiently powered to study rare diseases (e.g., oncology, sickle cell disease, etc) where other databases are too small to provide meaningful data.

The **Procedures Domain** contains data for all procedures performed on a patient enrolled in the PEDSnet database (>12,500 procedures captured). Similar to the diagnoses, procedures are recorded at each encounter and are mapped to standardized codes, which include SNOMED-CT, ICD-9 procedures, ICD-10 procedures, CPT-4, and HCPCS codes. Procedures are a useful measure for researchers and are often used in conjunction with diagnostic codes to improve case-finding accuracy and provide more detailed clinical data. It can also be used as a proxy for disease severity or clinical indications.

The PEDSnet database differentiates between the **Prescribed Medications Domain** and the **Dispensed Medications Domain**. Both domains capture recorded medications for a given encounter and map to a standardized RxNorm code. However, the dispensed medications domain provides more accurate information about the drugs that the patient actually received and can therefore more accurately assess a patient's medication record and disease control status. Both domains are vital in the conduct of epidemiologic research and are often used to construct a cohort of interest, assess disease status, define a patient's clinical course and health behaviors, and measure outcomes of interest.

The **Laboratory Test Results Domain** contains the orders and results for more than 400 laboratory tests. Examples include key components of a lipid panel, complete metabolic panel, complete blood counts, microbiologic cultures, liver function tests, urinalyses, viral panels, etc. This domain identifies more precise clinical outcomes and health determinants for patients. Researchers can evaluate and define disease progression more easily and measure and define disease status, severity, and control.

The **Visit Payer Domain** provides plan class (private/commercial, Medicaid/sCHIP, Medicare, other public, self-pay, other/unknown) and plan type (HMO, PPO, POS, fee-for-service, other/unknown) for every unique visit in the database. Researchers can use this domain to study the effect of payer visit type with health (and other) outcomes of interest. Data in this domain can also be used during statistical analyses as confounders, mediators, or effect modifiers to the association of interest.