

Welcome to Arcus

Arcus is a platform for CHOP researchers to access data, run analysis tools, and collaborate.

Getting Started

Log in to Arcus to start using it now. You will need to be a CHOP employee or a research collaborator with CHOP credentials, and [complete CITI training](#). After your CITI training has been registered with Arcus, you will need to log in and agree to the Arcus [Terms of Use](#).

Science-y Picture
or Research Data LifeCycle?

About Arcus

Privacy &
Security

FAQ

Available
Datasets

Data
Exploration
Tools

Analysis
Tools

Archived
Projects

Publications

Contact

QuickStart
Guide

Community
Forum

Research
Planning and
Management

Available Datasets

Clinical

ADR

The Arcus Data
Repository

# of Tables	54
-------------	----

ADR is data modeled from CHOP clinical data by Arcus to meet common research needs. There are several versions of the dataset (identified, deidentified, limited, and protected) used to prepare cohorts for Arcus scientific projects based on privacy requirements. To request data from the ADR, [Submit a Request For A New Scientific Project](#). If you already have a scientific project, [submit a data request here](#).

[Browse the ADR](#)[ADR Exploration Tools](#)

PEDSnet

Size	3.1 TB
------	--------

Cohort Size	1685558
-------------	---------

Total Number of Files	213
-----------------------------	-----

The PEDSnet database is part of our Arcus Archives Collections and captures clinical information for over 2.6 million pediatric patients, captured to support big data pediatric clinical research at participating institutions. This collection contains the Children's Hospital of Philadelphia (CHOP) only subset of the database. The relational database was exported to individual Comma Separated Value (CSV) tables that will be imported to BigQuery for research use.

[Browse PEDSnet](#)

Available Datasets

Helix

# of Tables	13788
-------------	-------

If the ADR does not contain the required CHOP clinical data for your research, researchers can also receive data from the Helix Data Warehouse for their Arcus Scientific Project. Helix is a cloud-based, modern, scalable, enterprise data analytics platform which supports all of CHOP's strategic data assets. Once you submit a Scientific Project Request, our team can assist in defining your cohort and determining which files you need to carry out your project. Please note if it is determined that your project involves a complex cohort, our team will refer you to work with the Clinical Reporting Unit.

[Browse Helix](#)

Clarity

# of Tables	27794
-------------	-------

If the ADR or Helix does not contain the CHOP clinical data you need for your project, researchers can receive data from The Clarity Data Warehouse for their project. Clarity is Epic's principal tool for managing clinical reporting data and is the primary data source for Epic's preferred analytical reporting tool. Clarity extracts data from the Chronicles database and stores it in a relational database on a separate, dedicated reporting server: the analytical database server, or Clarity server. Data that is normally stored in Chronicles items (or Caché globals) is structured in tables and columns on the Clarity database. Arcus can assist in identifying the data you need and accessing the appropriate data for your research. To get started Submit a Scientific Project Request. Please note if it is determined that your project involves a complex cohort, our team will refer you to work with the Clinical Reporting Unit.

[Browse Clarity](#)

Available Datasets

Omics

Division of Genomic Diagnostics

Size	213.7 TB
Cohort Size	13437
Total Number of Files	43823

The Arcus Division of Genomic Diagnostics (DGD) collection holds both raw and variant data for Whole Exome Sequencing (WES) and Whole Genome Sequencing (WGS). Certain individuals in the collection also have Children's Hospital of Philadelphia (CHOP) Rapid Targeted Analysis of the Genome for Infants (rTAG-I) Whole Genome Sequencing (WGS) results from the Baby Eagle program, which provides genetic testing for some of the smallest and sickest babies in the neonatal Intensive Care Unit (ICU), while gaining a better understanding of genome analysis.

[Browse DGD](#)[Variant Browser Tool](#)

The Birth Defects Biorepository (BDB)

Size	109.6 TB
Cohort Size	2516
Total Number of Files	18385

The purpose of the Birth Defects Biorepository (BDB) is to utilize the Children's Hospital of Philadelphia (CHOP) Center for Fetal Diagnosis and Treatment and Special Delivery Unit (SDU) and its well-developed facilities for high quality sample procurement and storage to create a platform for discovery of patient-specific variables (in utero, genetic and environmental) and disease-specific variables that underlie congenital defects and influence outcomes for patients, beginning during prenatal life and continuing across the lifespan.

[Browse BDB](#)[Variant Browser Tool](#)

Available Datasets

Surveys

Behavioral Health
Screening-Emergency
Department (BHS-ED)
Adolescent Data
Repository

Size	1.8 GB
Cohort Size	46913
Total Number of Files	32250

BHS-ED is data repository housing data on adolescent patients ages 14-18. The repository combines information from Epic and a non-Children's Hospital of Philadelphia (CHOP) electronic database, owned and managed by Medical Decision Logix (MDLogix), Inc. Currently, when patients in the Emergency Department (ED) complete a Behavioral Health Screen (BHS), MDLogix archives these responses in a cloud-based database and sends a Portable Document Format (PDF) of the BHS responses and results to the patient's Epic record. The Arcus team exports Fast Healthcare Interoperability Resources (FHIR) feed of this data from MDLogix. Data is available in either PDF or JavaScript Object Notation (JSON) FHIR format. Data collection in this repository is ongoing and currently consists of more than 20,000 unique patient records (of BHS survey results and corresponding medical record information) since 2013.

Browse BHS-ED