The CPI is designed and developed within the National Cancer Institute (NCI) Chilhood Cancer Data Initiative (CCDI) program, and is intended to provide a centralized index of participant data across multiple institutions.

The goal of the CPI is to manage and share multiple cross-linked participant IDs that represent the same individual by connecting various participant IDs from different studies/research institutions (domains). These mappings will empower researchers to explore complex questions, gain deeper insights into diseases, develop innovative therapies, and enhance existing treatments.

Identifiers and Domains

Participant ID: Public identifier that appears in a research dataset accessible to researchers (e.g., Kids First ID or GENIE ID). This also includes global identifiers used more broadly (e.g., COG USI).

Domain: Groups of unique participant IDs, Domains can be categorized as:

- **Institutional Identifier**. Typically refers to a larger project/network namespace of unique participant identifiers. Current institutional identifiers include GENIE (AACR Project GENIE), GMKF (Gabriella Miller Kids First), PCDC (Pediatric Cancer Data Commons), StJude_CompBIOID (St Jude Comp Bio IDs), Treehouse (Treehouse Childhood Cancer Initiative and USI (COG USI).
- **Dataset.** Refers to a single dataset for a particular study where data files for the participants can be found. This is typically a dbGap phs accession number – for example phs000720 (dbGaP accession: Genomic sequencing of Pediatric Rhabdomyosarcoma).
- **Study.** Represents programmatic references to research initiatives, such as TARGET, or institution's internal study synonyms for dataset domains such as SD_AQ9KVN5P, equivalent to phs002276.

The CPI is envisioned as an API-mediated service with a minimal web application to serve as an administrative user interface. CPI will not hold PII or any participant data—just IDs. Access to data with which those IDs are associated will remain under control of primary data source owners.

Access and Support

The CPI API will be made available to applications and services authorized by NCI. Integration with the CPI is currently under development for certain CCDI applications; other interested system owners may initiate a request for access by emailing NCIChildhoodCancerDataInitiative@mail.nih.gov.