



CHILDHOOD CANCER CLINICAL DATA COMMONS (C3DC) User Guide

03/05/2025

Version	Date	Description	Author
1.0.0	03/05/2024	Initial Release	C3DC Team
2.0.0	05/01/2024	Release 2	C3DC Team
3.0.0	07/31/2024	Release 3	C3DC Team
4.0.0	11/07/2024	Release 4	C3DC Team
5.00	03/05/2025	Release 5 (version 1.4.0)	C3DC Team

Table of Contents

INTRODUCTION AND OVERVIEW	4
C3DC HOME PAGE	5
C3DC EXPLORE PAGE	7
FACET SEARCH	8
CHART VISUALIZATION	8
TABLE VISUALIZATION	9
DOWNLOAD HARMONIZED DATA	9
COHORT SELECTOR	10
C3DC COHORT ANALYZER (NEW KEY FEATURE)	12
C3DC STUDIES PAGE	19
C3DC STUDIES DETAILS PAGE	19
C3DC DATA MODEL PAGE	21
C3DC RESOURCE PAGE	22
C3DC ABOUT PAGE	23
ANNOUNCEMENTS	24
RELEASE NOTES	25
USER GUIDE	26

Introduction and Overview

The Childhood Cancer Clinical Data Commons (C3DC) enables searching demographic and phenotypic clinical data of childhood cancers. These data have been harmonized to a standard set of common data elements (CDEs). C3DC empowers researchers to search for participant-level data to create synthetic cohorts and export data for analysis.

This document describes a high-level overview of the features of C3DC. Investigators are encouraged to explore C3DC themselves, using this guide as a primer.

C3DC Home Page

The Home Page allows users to navigate through key sections, including Explore, Cohort Analyzer, Data Model, and About pages.

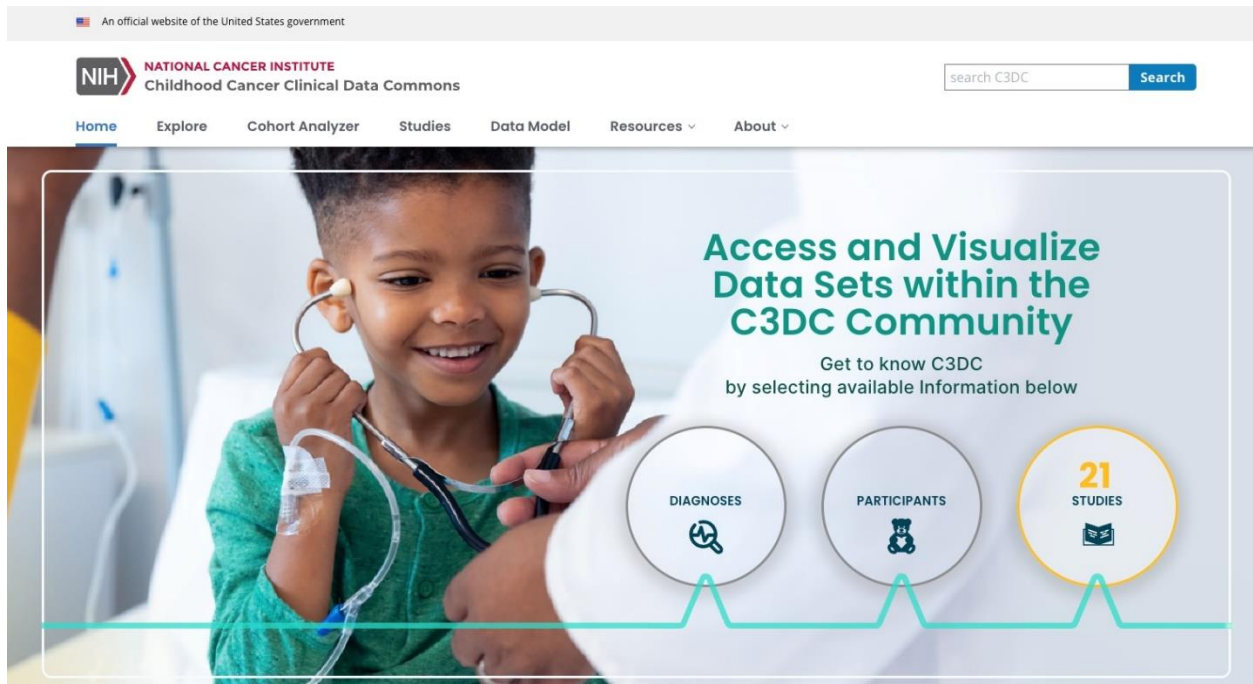


Figure 1: C3DC Home Page (top)

At the bottom of the Home page, there are links to brief descriptions of key sections of the C3DC application.

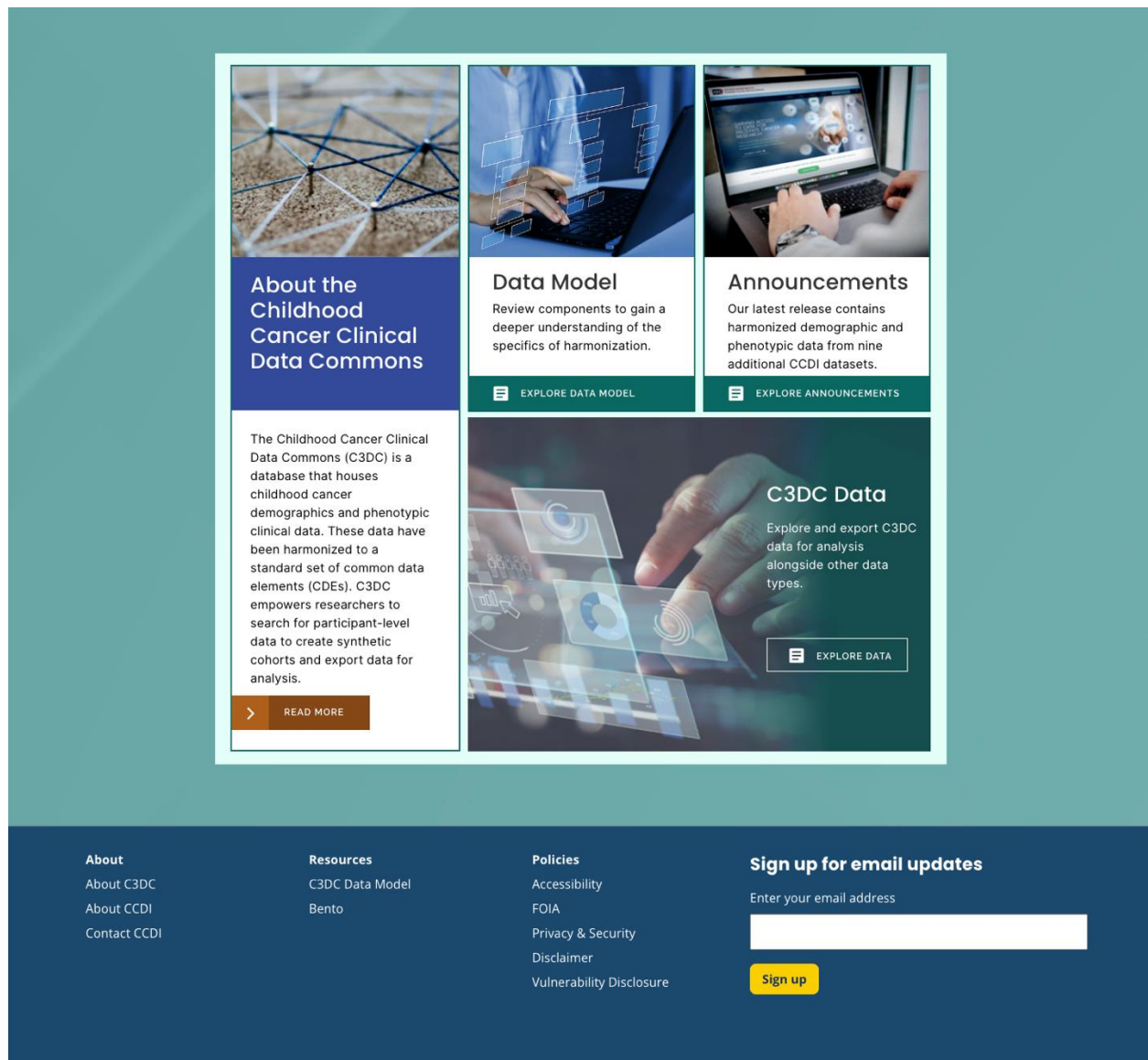


Figure 2: C3DC Home Page (bottom)

C3DC Explore Page

The Explore Page is the main interface for searching and visualizing data. Users can refine results using faceted search options and analyze participant-level details.

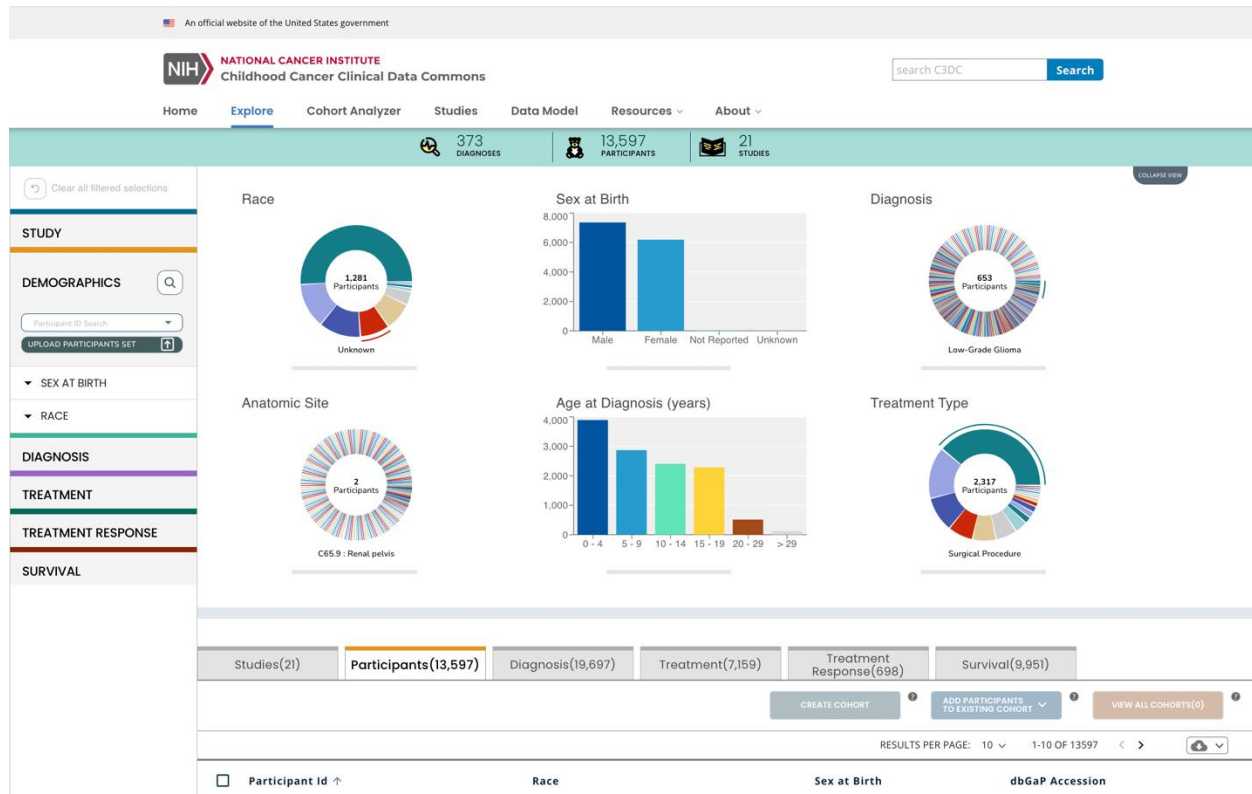


Figure 3: The Explore Page – Faceted Search (left), Visualization (top), and Data Tables (bottom)

Facet Search: Users can select a subset of the childhood cancer participants by choosing filtering options from drop-down lists within six main categories: Study, Demographics, Diagnosis, Treatment, Treatment Response, and Survival.

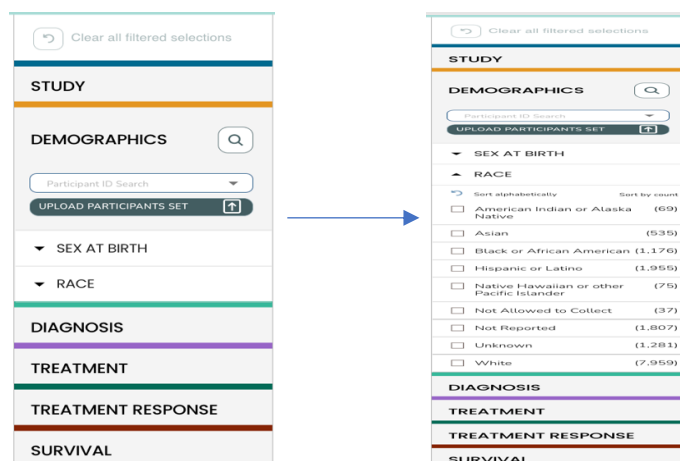


Figure 4: Performing a faceted search on the C3DC Explore Page

Chart Visualization: The **Stats Bar**, **Visualization Section**, and **Data Table** dynamically update based on applied filters.

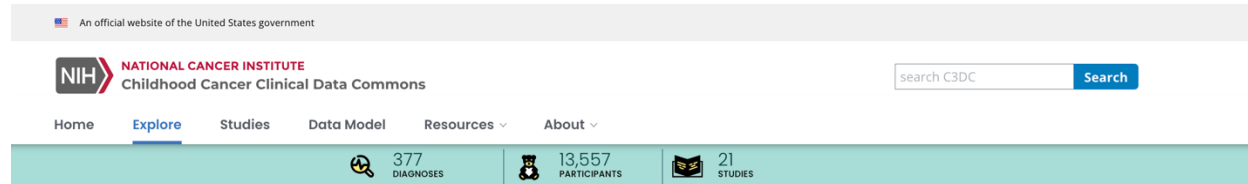


Figure 5: Results returned from a faceted search in the Stats bar

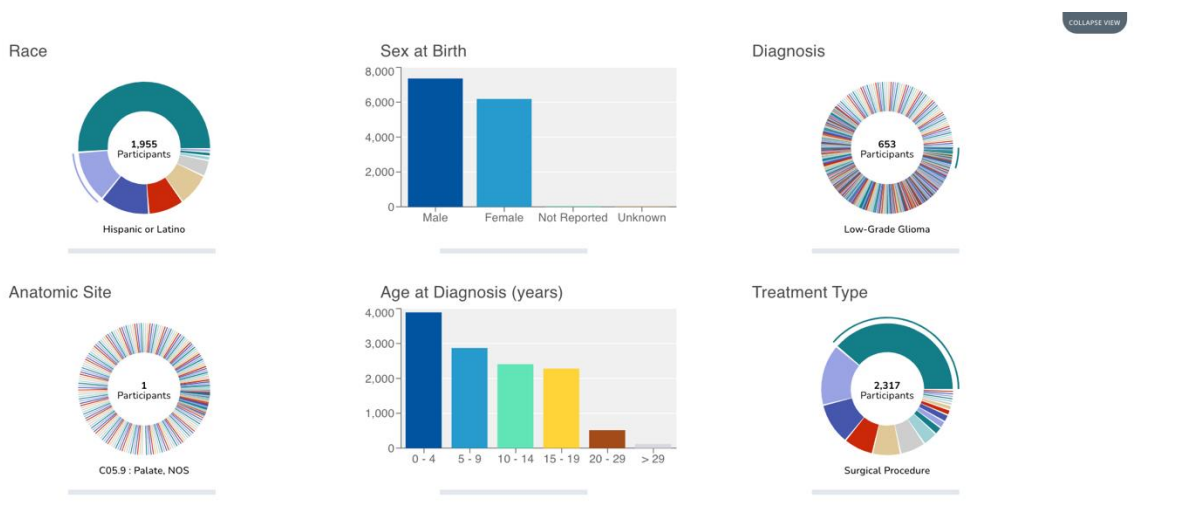


Figure 6: Results returned from a faceted search in six graphs in the Visualization section

Table Visualization: Displays participant details with **tooltips explaining table headers**.

Count of Diagnosis Record					
Studies(21)	Participants(13,597)	Diagnosis(19,697)	Treatment(7,159)	Treatment Response(698)	Survival(9,951)
CREATE COHORT [?] ADD PARTICIPANTS TO EXISTING COHORT [?] VIEW ALL COHORTS(2) [?]					
RESULTS PER PAGE: 10 1-10 OF 13597					
1 row(s) selected					
Participant Id ↑	Race	Sex at Birth	dbGaP Accession		
<input checked="" type="checkbox"/> 00301d78915737fa100f	White	Female	phs002431		
<input type="checkbox"/> 0061cbb0846973206fcf	White	Male	phs002431		
<input type="checkbox"/> 0065af91e89ee2859595	Hispanic or Latino;White	Female	phs002431		
<input type="checkbox"/> 008b04a84717e7d007a4	White	Male	phs002431		
<input type="checkbox"/> 009f9eafdc9602a28fb	White	Female	phs002431		
<input type="checkbox"/> 00c5e4372375eb3627f3	White	Female	phs002431		
<input type="checkbox"/> 00dfd38d509984eada22	White	Male	phs002431		
<input type="checkbox"/> 00e3d1d383c8c08a25d2	White	Male	phs002431		
<input type="checkbox"/> 00eff50ac98a200da9f9	White	Male	phs002431		
<input type="checkbox"/> 01504e3b7031a9ec2b2	White	Female	phs002431		

Figure 7: Results returned from a faceted search in the Table section

Download Harmonized Data: Users can download the contents of the Studies, Participants, Diagnosis, Treatment, Treatment Response, and Survival tabs by selecting the "Download Data" button under the table tab headers. Users can download filtered data in **CSV or JSON formats**.

A1	Participant Id	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Participant Id	Diagnosis ID	Diagnosis	Diagnosis Cl	Diagnosis Ba	Diagnosis Co	Disease Phat	Tumor Classi	Anatomic Sit	Age at Diagn	Toronto Chilt	Tumor Grade	Tumor Stage	Tumor Stage	Study ID	Dbgap Accession	
2	PBCHDL	02296835-e61b-4b87-a9fa-1d105b3977bc	9421/1: Pilot	ICD-O-3.2	Clinical	Not Reported	Initial Diagn	Primary	C71.6: Cereb	2005	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002790.v	phs002790
3	PT_17H2AASA	0743006ecc41e772cad3655f6be43221_1	High-Grade G	CNS Diagnos	Pathological	High-grade gli	Initial Diagn	Primary	C71.6: Cereb	2712	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
4	PT_RC1THQJQ	0bee93a8e1635b5853ba0536288c5d3_1	9400/0: Astro	ICD-O-3.2	Pathological	Low-grade gli	Initial Diagn	Primary	C71.6: Cereb	1548	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
5	PT_V9WP047R	0e958c55474d5ee5508a344efc825e7_1	Medulloblast	CNS Diagnos	Pathological	Medulloblast	Initial Diagn	Primary	C71.6: Cereb	2212	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
6	PT_D286YN6	0e9ad0f3e7b9c003aac096f9ca87513_1	9400/0: Astro	ICD-O-3.2	Pathological	Low-grade gli	Initial Diagn	Primary	C71.6: Cereb	1805	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
7	PT_8N82561	0f04dbdc3013b287ca76698578798026_5	9508/3: Asypl	ICD-O-3.2	Pathological	Atypical Tera	Post-Mortem	Primary	C71.6: Cereb	3562	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
8	PT_962TCBVR	0f0c640db8a4c55389e64845682c5c7a_1	Low-Grade Gl	CNS Diagnos	Pathological	Low-grade gli	Progression	Primary	C71.6: Cereb	2402	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
9	PT_GAKHC0Y2	1466dd5ca7bfb5e937172709cb03db_1	Low-Grade Gl	CNS Diagnos	Pathological	Low-grade gli	Initial Diagn	Primary	C71.6: Cereb	1372	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
10	PT_6VKN63HK	1628e90f2c14d37f5606a5b3dea3f734_1	9400/0: Astro	ICD-O-3.2	Pathological	Low-grade gli	Initial Diagn	Primary	C71.6: Cereb	3698	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
11	PT_64JHC2BH	1b01de1a7be9b342681d2c9a19467eb0_1	Low-Grade Gl	CNS Diagnos	Pathological	Low-grade gli	Initial Diagn	Primary	C71.6: Cereb	1473	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
12	PT_KJKY2JZ	1b555027a5cb4a089f685f70602e2b4_1	9400/0: Astro	ICD-O-3.2	Pathological	Low-grade gli	Initial Diagn	Primary	C71.6: Cereb	2326	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
13	PT_962TCBVR	1b5dbb2568ddb37deb896fa1a6c9ac25_1	Low-Grade Gl	CNS Diagnos	Pathological	Low-grade gli	Recurrent D	Primary	C71.6: Cereb	1717	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	phs002517
14	PBBJPZ	1f9cb1f9-fac8-4661-a060-0b79274000d5	9421/1: Pilot	ICD-O-3.2	Clinical	Not Reported	Initial Diagn	Primary	C71.6: Cereb	1040	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002790.v	phs002790

Figure 8: Download Diagnosis Table tab as CSV file format

Cohort Selector: The Cohort Selector enables users to create and manage up to 20 cohorts. This feature offers flexibility to researchers, allowing them to create cohort groups according to their specific requirements.

Studies(21)

Participants(13,597)

Diagnosis(19,697)

Treatment(7,159)

Treatment Response(698)

Survival(9,951)

CREATE COHORT

ADD PARTICIPANTS TO EXISTING COHORT

VIEW ALL COHORTS(2)

RESULTS PER PAGE: 10 1-10 OF 13597

1 row(s) selected

Participant Id ↑	Race	Sex at Birth	dbGaP Accession
<input checked="" type="checkbox"/> 00301d78915737fa100f	White	Female	phs002431
<input type="checkbox"/> 0061cbb0846973206fcf	White	Male	phs002431
<input type="checkbox"/> 0065af91e89ee2859595	Hispanic or Latino;White	Female	phs002431
<input type="checkbox"/> 008b04a84717e7d007a4	White	Male	phs002431
<input type="checkbox"/> 009f9efaedc9602a28fb	White	Female	phs002431
<input type="checkbox"/> 00c5e4372375eb3627f3	White	Female	phs002431
<input type="checkbox"/> 00dfd38d509984eada22	White	Male	phs002431
<input type="checkbox"/> 00e3d1d383c8c08a25d2	White	Male	phs002431
<input type="checkbox"/> 00eff50ac98a200da9f9	White	Male	phs002431
<input type="checkbox"/> 01504e3b7031a9eec2b2	White	Female	phs002431

RESULTS PER PAGE: 10 1-10 OF 13597

Figure 9: Cohort Selection features visible on the Explore page Table tabs (see choice buttons in orange rectangle)

Users can do the following:

- **Create Cohort:** Select participant IDs from the facet search and create a new cohort. Users can name and describe the cohort for easy reference.
- **Add Participants to Existing Cohort:** Add participants to existing cohorts or remove them. Entire cohorts can also be deleted as needed.
- **View All Cohort(s):** View a list of all created cohorts, making it easier to manage and analyze groups.
 - **Cohort ID:** Create your own ID to identify saved cohorts
 - **Cohort Description:** Create a description for saved cohorts
 - **Save Changes:** Save the changes made to the selected cohort. This includes changes to cohort ID, cohort description, and any participants removed.
 - **Download Selected Cohort:**
 - Download the selected cohort in one of two formats.
 - **Manifest CSV:** a list of participant IDs and high-level data.
 - **Metadata JSON:** a JSON file containing all metadata information for the participants in the selected cohort.

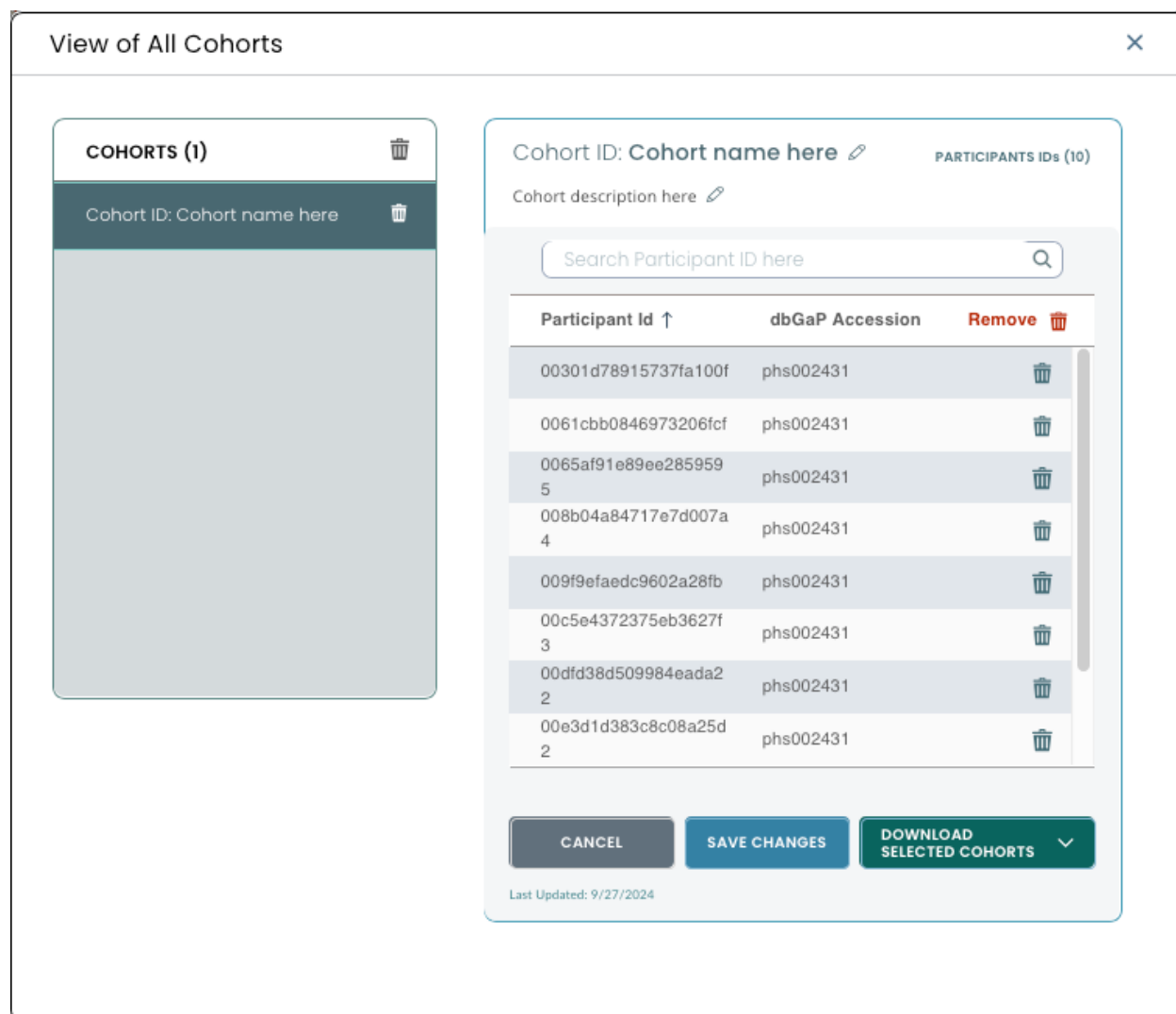


Figure 10: View All Cohorts and edit capabilities

C3DC Cohort Analyzer (New Key Feature)

The Cohort Analyzer is designed to compare up to three cohorts and visualize their intersections through an interactive Venn diagram and a corresponding data table. This feature leverages cohorts created on the Explore page, enabling users to analyze key relationships and distinctions based on unique values between datasets effectively.

Customizable Properties:

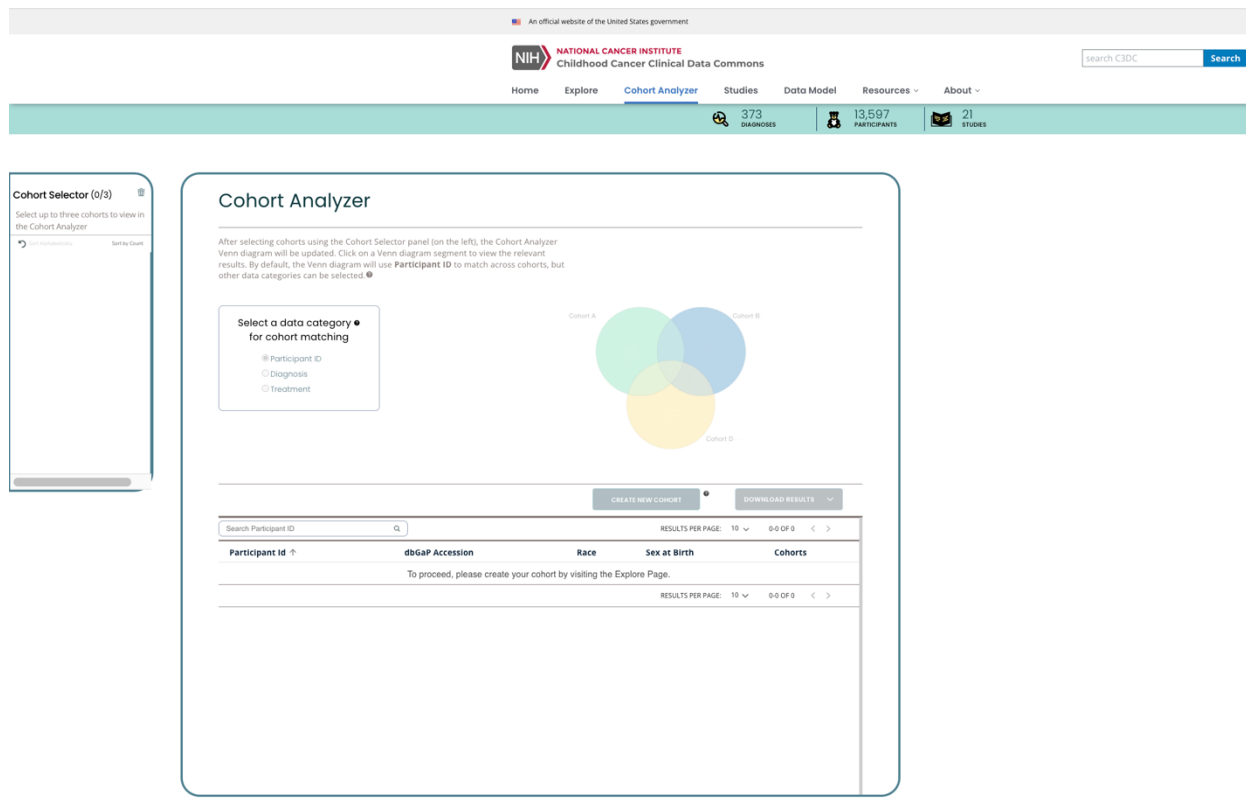
The radio buttons allow users to select more than one property for comparison. Available properties include:

- Participant ID
- Diagnosis
- Treatment

Enhanced Analytical Capabilities:

Users will be able to visualize overlaps and unique attributes within each cohort. In addition, users can:

- Investigate specific sections of the Venn diagram to view participant-level details from the corresponding table view
- Export results, including the data table and Venn diagram, for further analysis or integration into other platforms.
- Use advanced filters to refine cohort comparisons, such as narrowing by treatment or specific diagnosis.



To start using the Cohort Analyzer, you will first need to select the cohorts you want to analyze. As you add cohorts, the system will automatically keep track of your cohorts on the left side Cohort Selector. This tool's functionality adapts based on the number of selected cohorts, ensuring a customized analysis.

Select your first cohort by clicking the check box in the Cohort Selector sidebar. The Venn diagram and table will update to display the cohort information based on the participant ID, diagnosis or treatment based on the radio button selection. In this example, we are using the participant ID, thus the table will contain properties specific to the participant as well as show to what cohorts the participant belongs.

An official website of the United States government

NIH NATIONAL CANCER INSTITUTE
Childhood Cancer Clinical Data Commons

search C3DC Search

Home Explore Cohort Analyzer Studies Data Model Resources About

373 DIAGNOSES 13,597 PARTICIPANTS 21 STUDIES

Cohort Selector (1/3)

Select up to three cohorts to view in the Cohort Analyzer

Search Cohorts

☒ Cohort 1 (12)

☐ Cohort 10 (11)

☐ Cohort 11 (10)

☐ Cohort 2 (13)

☐ Cohort 3 (13)

☐ Cohort 4 (11)

☐ Cohort 5 (15)

☐ Cohort 6 (15)

☐ Cohort 7 (16)

☐ Cohort 8 (17)

☐ Cohort 9 (18)

Cohort Analyzer

After selecting cohorts using the Cohort Selector panel (on the left), the Cohort Analyzer Venn diagram will be updated. Click on a Venn diagram segment to view the relevant results. By default, the Venn diagram will use **Participant ID** to match across cohorts, but other data categories can be selected.

Select a data category for cohort matching

☒ Participant ID

☐ Diagnosis

☐ Treatment

Cohort 1 (12)

12

CREATE NEW COHORT

DOWNLOAD RESULTS

Search Participant ID

RESULTS PER PAGE: 10 1-10 OF 12

Participant ID	dbGaP Accession	Race	Sex at Birth	Cohorts
PBCKIZ	phs002790	Hispanic or Latino	Female	<input type="checkbox"/>
PBCKPG	phs002790	Not Reported	Female	<input type="checkbox"/>
PBCKWK	phs002790	White	Male	<input type="checkbox"/>
PBCMUI	phs002790	Not Reported	Female	<input type="checkbox"/>
R1403703	phs002518	Hispanic or Latino	Male	<input type="checkbox"/>
R12325935	phs002518	White	Male	<input type="checkbox"/>
R19731467	phs002518	Asian	Female	<input type="checkbox"/>
R20517902	phs002518	Hispanic or Latino;White	Male	<input type="checkbox"/>
R23374497	phs002518	White	Female	<input type="checkbox"/>
R24320130	phs002518	Unknown	Male	<input type="checkbox"/>

RESULTS PER PAGE: 10 1-10 OF 12

Figure 12: One Cohort Selected: You have created and selected one cohort.

Select another cohort in the Cohort Selector to see the Venn diagram and table update again. This time, if there are common participants between both cohorts, the diagram will show the shared participants in the intersection between the two. In the table, with none of the Venn diagram selected, it will display all participants and their respective cohort.

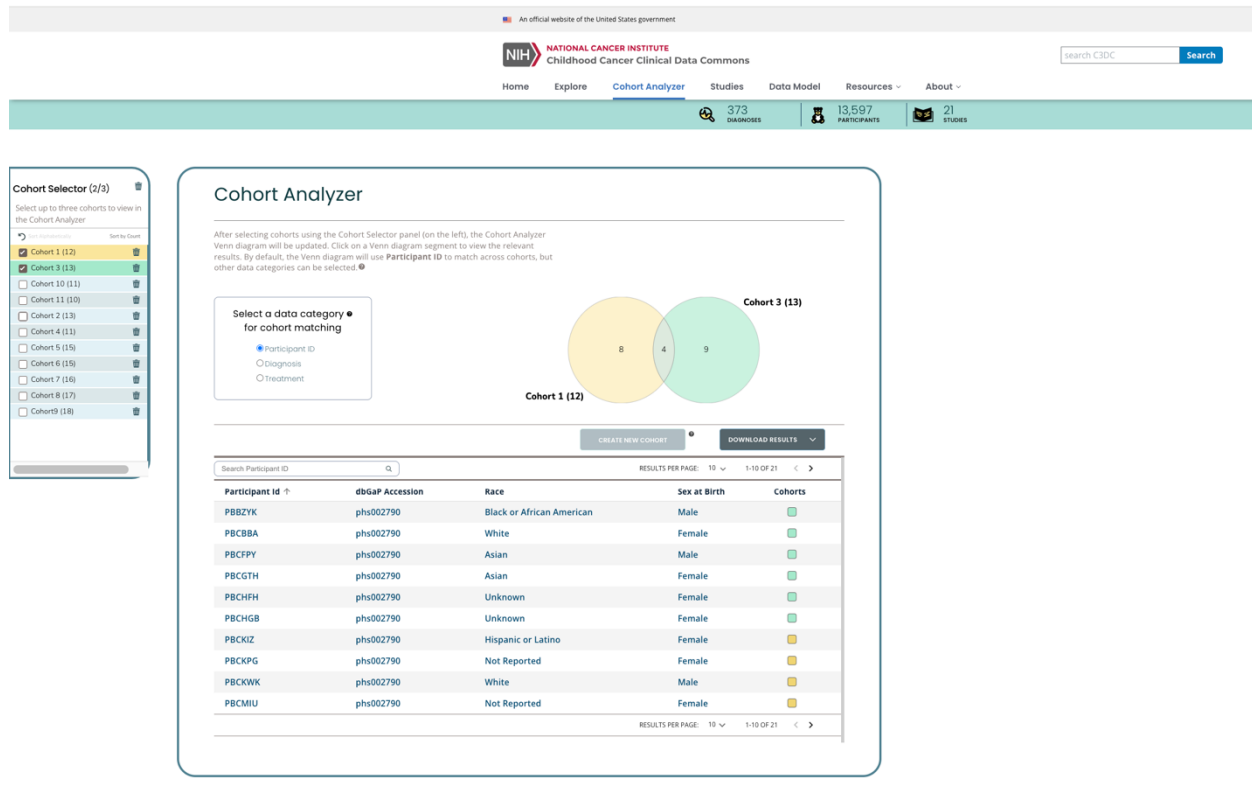


Figure 13: Two Cohorts Selected: You have selected two cohorts. Visualize shared and unique data points between these cohorts.

Select a third and final cohort. The Venn diagram and table will update once more with all the participant level data.

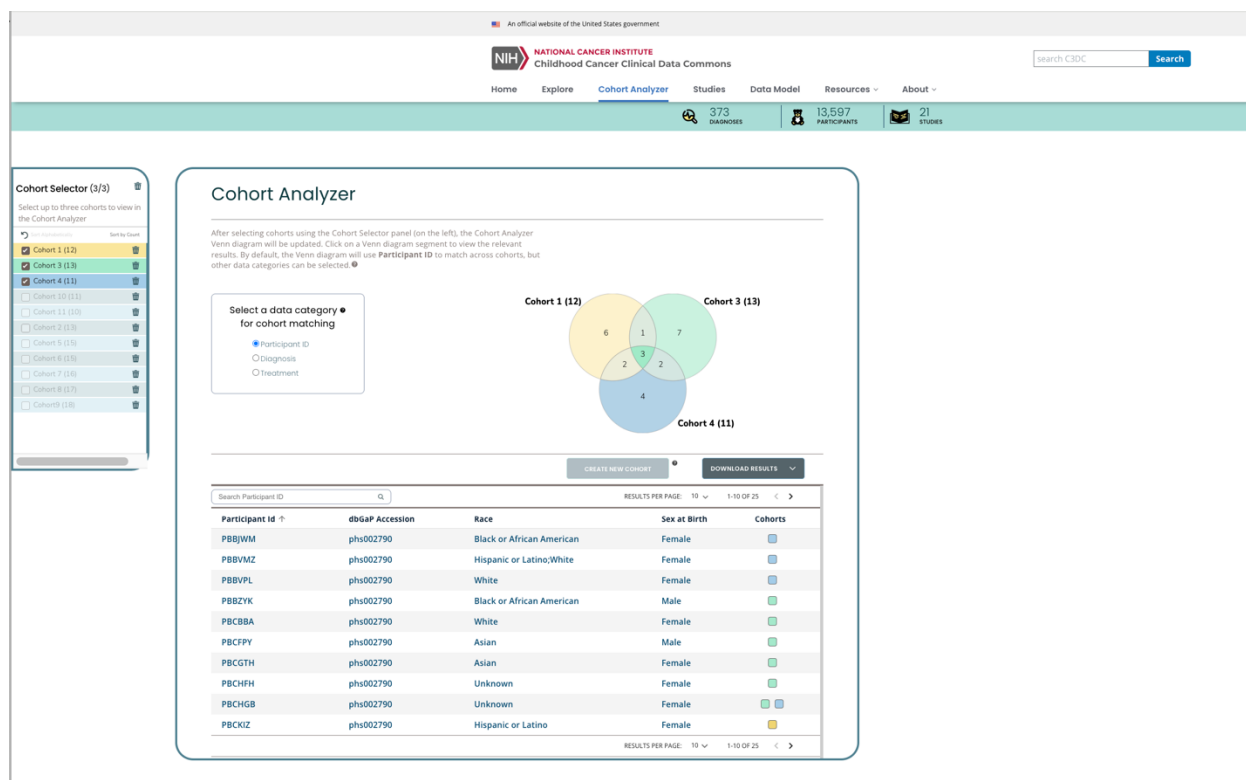


Figure 14: Three Cohorts Selected: You have selected three cohorts. Explore their intersections and unique attributes using the Venn diagram.

Please note that the number in parentheses by the cohort's name in the Venn diagram represents the count of unique records for that radio button selection. The number inside the Venn diagram sections are the count of unique values for that radio button selection. Finally, the count next to your cohort in the Cohort Selection side bar indicates the total participants in your cohort.

NATIONAL CANCER INSTITUTE

Childhood Cancer Clinical Data Commons

Home

Explore

Cohort Analyzer

Studies

Data Model

Resources

About

373

DIAGNOSES

13,597

PARTICIPANTS

21

STUDIES

Search

Cohort Selector (3/3)

Select up to three cohorts to view in the Cohort Analyzer

Cohort 1 (12)

Cohort 3 (13)

Cohort 4 (11)

☐ Cohort 2 (10)
 ☐ Cohort 11 (10)
 ☐ Cohort 2 (13)
 ☐ Cohort 5 (15)
 ☐ Cohort 6 (15)
 ☐ Cohort 7 (16)
 ☐ Cohort 8 (17)
 ☐ Cohort9 (18)

Select a data category for cohort matching

☒ Participant ID
 ☐ Diagnosis
 ☐ Treatment

Cohort 1 (12)

Cohort 3 (13)

Cohort 4 (11)

CREATE NEW COHORT

DOWNLOAD RESULTS

Search Participant ID

Q

RESULTS PER PAGE: 10

1-3 OF 3

<

>

Participant id	dbGaP Accession	Race	Sex at Birth	Cohorts
R1403703	phs002518	Hispanic or Latino	Male	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
R20517902	phs002518	Hispanic or Latino/White	Male	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
R23374497	phs002518	White	Female	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

RESULTS PER PAGE: 10

1-3 OF 3

<

>

17

The user will see the intersections of all three cohort. Additionally, the user will also see intersections between two cohorts. Clicking on the desired intersection will result in the table being updated accordingly with metadata for those selected participants.

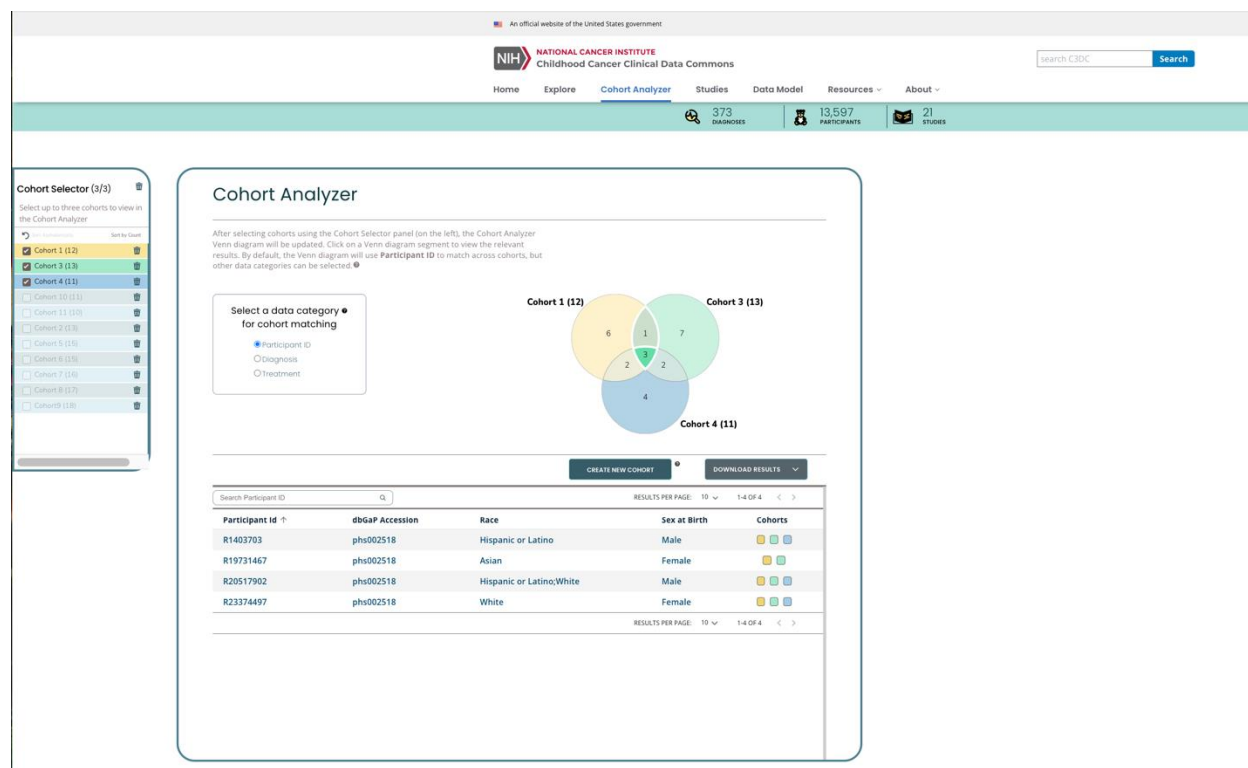


Figure 16: View the specific intersections between selected Cohorts

The Cohort analyzer offers a powerful method to explore how various clinical attributes overlap and differ across multiple groups. By visualizing the shared and unique data points using a Venn diagram, you can identify common patterns or variations in key clinical variables such as diagnosis, treatment, and participant characteristics. This analysis helps reveal underlying trends in the clinical data that may be crucial for research, such as identifying which treatment protocols are common across cohorts or exploring the presence of specific diagnoses.

C3DC Studies Page

Users can navigate to the Studies Page to view the list of dbGaP accessions, study names, and counts for participants and diagnoses.

dbGaP Accession ↑	Study Name	Participants Count	Diagnosis Count
phs000466	TARGET: Kidney, Clear Cell Sarcoma of the Kidney (CCSK)	13	1
phs000467	TARGET: Neuroblastoma (NBL)	1,119	3
phs000470	TARGET: Kidney, Rhabdoid Tumor (RT)	69	1
phs000471	TARGET: Kidney, Wilms Tumor (WT)	652	1
phs000720	Genomic Sequencing of Pediatric Rhabdomyosarcoma	403	5
phs001437	Pediatric Preclinical Testing Consortium (PPTC)	267	28
phs002371	Human Tumor Atlas Network (HTAN)	30	4
phs002430	Genomic Analysis in Pediatric Malignancies	231	52
phs002431	Enhancement of Data Sharing in Pediatric, Adolescent and Young Adult Cancers	611	156
phs002504	UCSF Database for the Advancement of JMML - Integration of Metadata with "Omic" Data	195	1

Figure 17: The Studies Page

C3DC Studies Details Page: By clicking on the dbGaP accession number, users can access detailed information about the studies. This action will redirect users to the dbGaP page to view a high-level overview of the information. Currently, source data files are only available for open access data (TARGET datasets phs000466, phs000467, phs000468, phs000469, phs000470, and phs000471) and manifest metadata is available for all other CCDI studies. If you are interested in accessing the controlled access data, please follow this [link](#) for instructions on how to access it.

For other CCDI studies, source data can be found in the [CCDI Hub](#).

An official website of the United States government

NIH NATIONAL CANCER INSTITUTE
Childhood Cancer Clinical Data Commons

search C3DC Search

Home Explore Cohort Analyzer Studies Data Model Resources About

3 DIAGNOSES 1,119 PARTICIPANTS 1 STUDIES

Home > Studies > Study Code phs000467

C3DC Studies

DBGAP ACCESSION: [phs000467](#)

Overview

STUDY DESCRIPTION

There are ~214 fully characterized patient cases with neuroblastoma (all tumor/normal pairs, 10 with relapse sample as well) that will make up the TARGET NBL dataset, along with some cell lines and xenografts. The dataset includes 24 45 cases as well. Each fully characterized case has gene expression, tumor and paired normal copy number analyses, methylation and comprehensive next-generation sequencing to include whole genome and/or whole exome sequencing. A majority of these cases will also have mRNA-seq and methylation data available as well. There are additionally a large number of cases, both low and high risk, with partial molecular characterization to include some next generation and targeted Sanger sequencing making this a large and informative genomic dataset. Please visit the TARGET website, <https://www.cancer.gov/ccq/research/genome-sequencing/target>, for additional information on this and other TARGET genomics projects. **Please see the TARGET Publication Guidelines for updated details on the sharing of any TARGET substudy data.**

Source Files: [TARGET_NBL_ClinicalData_Discovery_20220125.xlsx](#)
[TARGET_NBL_ClinicalData_Validation_20220125.xlsx](#)

PARTICIPANTS
1,119

SURVIVAL RECORDS
1,119

DIAGNOSES
3

ANATOMICAL SITES
48

Figure 18A: The Studies Details page – downloadable Open Access Source file data for TARGET datasets

An official website of the United States government

NIH NATIONAL CANCER INSTITUTE
Childhood Cancer Clinical Data Commons

search C3DC Search

Home Explore Cohort Analyzer Studies Data Model Resources About

28 DIAGNOSES 267 PARTICIPANTS 1 STUDIES

Home > Studies > Study Code phs001437

C3DC Studies

DBGAP ACCESSION: [phs001437](#)

Overview

STUDY DESCRIPTION

The Pediatric Preclinical Testing Consortium (PPTC) is addressing the unmet need of streamlining the development of new therapies for childhood cancers. The PPTC seeks to develop robust biomarkers of anticancer drug activity, and the majority of these are predicted to be genetic mutations that can be detected in tumor DNA and/or RNA. In order to design the most impactful experiments that can be rapidly translated to the clinic, PPTC investigators require a complete genomic characterization of the patient-derived xenograft tumor models that are utilized across the consortium. This will not only allow for the most robust experimental design, but also will increase the engagement of industry partners who seek collaborators poised to provide the proof-of-concept necessary for drugs in their development pipelines. All data and models will be made available to academically qualified investigators.

Manifest Files: [phs001437_CCDI_Study_Manifest_v2.0.0.xlsx](#)

PARTICIPANTS
267

SURVIVAL RECORDS
86

DIAGNOSES
28

ANATOMICAL SITES
35

Figure 18B: The Studies Details page – downloadable manifest metadata files for all other CCDI studies

C3DC Data Model Page

The [data model](#) is developed collaboratively with multiple organizations to establish standard terms for pediatric cancer. In this harmonization effort, we are using CDEs (Common Data Elements) to enhance data accuracy, consistency, and interoperability across health research studies. CDEs are defined in the caDSR (Cancer Data Standards Registry and Repository) and provide controlled terms, vocabularies, detailed information on data representation, and robust metadata. The C3DC data model schema consists of well-defined classes with attributes and permissible values. The model focuses on a set of data elements, with plans to include more elements in future releases.

An official website of the United States government

NIH NATIONAL CANCER INSTITUTE
Childhood Cancer Clinical Data Commons

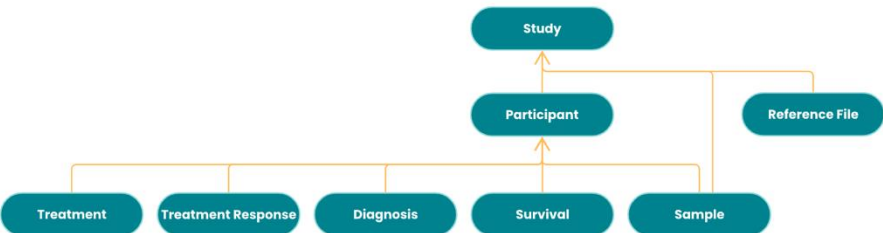

search C3DC Search

Home Explore Studies **Data Model** Resources About

C3DC Data Model

The C3DC harmonized [Data Model](#) unites multiple childhood cancer studies and drives the search experience. Initially, the C3DC database model will focus on a limited set of data elements, such as participant diagnosis, to establish processes for data ingestion, harmonization, data modeling, and storage. The C3DC model is developed collaboratively with contributions from Data for the Common Good at University of Chicago, Frederick National Laboratory, NCI's Semantics Infrastructure, Cancer Research Data Commons, and National Cancer Childhood Registry team members. The resulting model will collectively establish the standard terms for pediatric cancer and will serve as guidance to researchers

NCI encourages researchers to use CDEs to enhance data accuracy, consistency, and interoperability across diverse areas of health research. CDEs are pieces of data captured in multiple datasets across different studies. The caDSR, short for the [Cancer Data Standards Registry and Repository](#) is a structured repository for clinical and research data. Its content semantically defines data through controlled terms and vocabularies, offering detailed information about data representation, including permissible values, data types (numeric, date, text), and the ability to associate various local or source names with the same data value. Moreover, CDEs provide robust metadata, encompassing references to external standards, usage instructions, and examples of expected data formats. The C3DC target schema will consist of well-defined classes with attributes and permissible values.

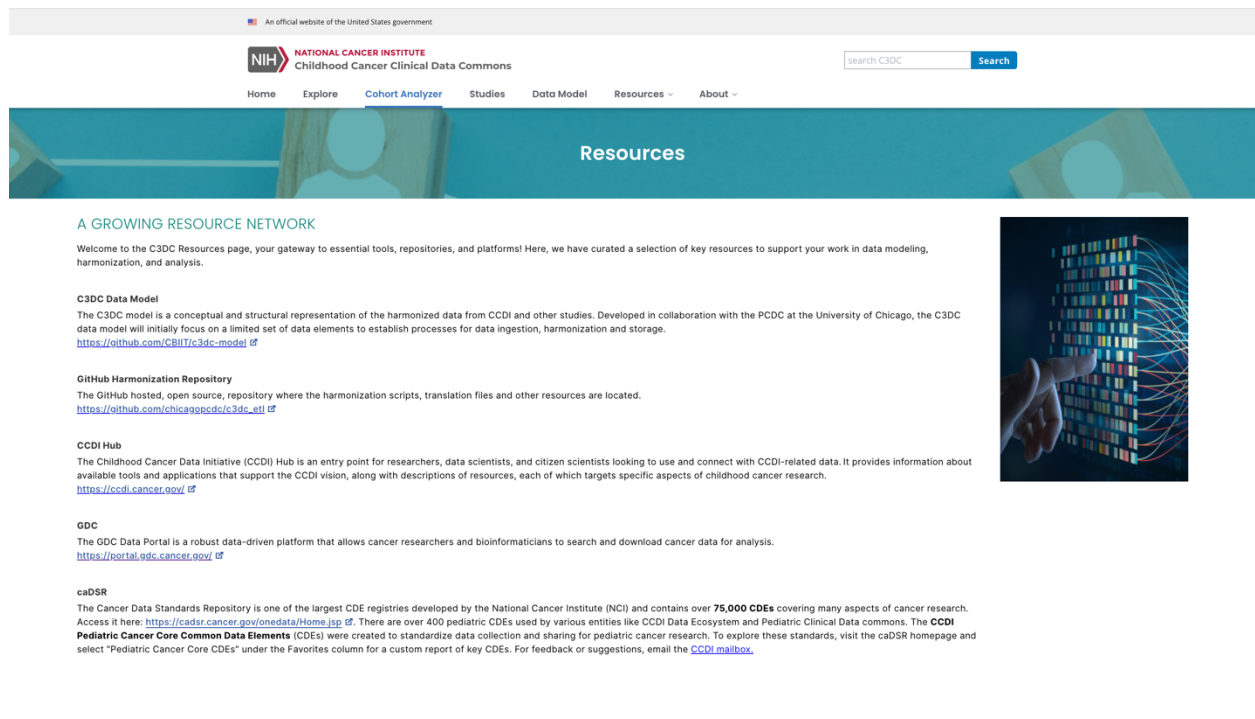


```
graph TD; Study --> Participant; Study --> ReferenceFile[Reference File]; Participant --> Treatment; Participant --> TreatmentResponse[Treatment Response]; Participant --> Diagnosis; Participant --> Survival; Participant --> Sample;
```

Figure 19: The Data Model page

C3DC Resource Page

By clicking on each resource, users can access a range of useful tools and information available on the site.



The screenshot displays the C3DC Resource Page, an official website of the United States government. The header features the NIH logo and the text "NATIONAL CANCER INSTITUTE Childhood Cancer Clinical Data Commons". A search bar is located in the top right corner. The main navigation menu includes links for Home, Explore, Cohort Analyzer, Studies, Data Model, Resources, and About. The page title is "Resources".

A GROWING RESOURCE NETWORK

Welcome to the C3DC Resources page, your gateway to essential tools, repositories, and platforms! Here, we have curated a selection of key resources to support your work in data modeling, harmonization, and analysis.

C3DC Data Model

The C3DC model is a conceptual and structural representation of the harmonized data from CCDI and other studies. Developed in collaboration with the PCDC at the University of Chicago, the C3DC data model will initially focus on a limited set of data elements to establish processes for data ingestion, harmonization and storage.

<https://github.com/CBILTric3dc-model>

GitHub Harmonization Repository

The GitHub hosted, open source, repository where the harmonization scripts, translation files and other resources are located.

https://github.com/chicago-cdc/c3dc_etl

CCDI Hub

The Childhood Cancer Data Initiative (CCDI) Hub is an entry point for researchers, data scientists, and citizen scientists looking to use and connect with CCDI-related data. It provides information about available tools and applications that support the CCDI vision, along with descriptions of resources, each of which targets specific aspects of childhood cancer research.

<https://ccdi.cancer.gov/>

GDC

The GDC Data Portal is a robust data-driven platform that allows cancer researchers and bioinformaticians to search and download cancer data for analysis.

<https://portal.gdc.cancer.gov/>

caDSR

The Cancer Data Standards Repository is one of the largest CDE registries developed by the National Cancer Institute (NCI) and contains over **75,000 CDEs** covering many aspects of cancer research. Access it here: <https://caDSR.cancer.gov/onedata/Home.jsp>. There are over 400 pediatric CDEs used by various entities like CCDI Data Ecosystem and Pediatric Clinical Data commons. The **CCDI Pediatric Cancer Core Common Data Elements** (CDEs) were created to standardize data collection and sharing for pediatric cancer research. To explore these standards, visit the caDSR homepage and select "Pediatric Cancer Core CDEs" under the Favorites column for a custom report of key CDEs. For feedback or suggestions, email the [CCDI mailbox](#).




Figure 20: The Resource page

C3DC About Page

Users can navigate to the About Page by clicking the link on the Home Page menu bar, where you will find more information about the content of C3DC. This includes details such as dataset and data model information as well as links to useful resources. There are dropdown menus to access the following documents:

- Announcements
- Release Notes
- User Guide

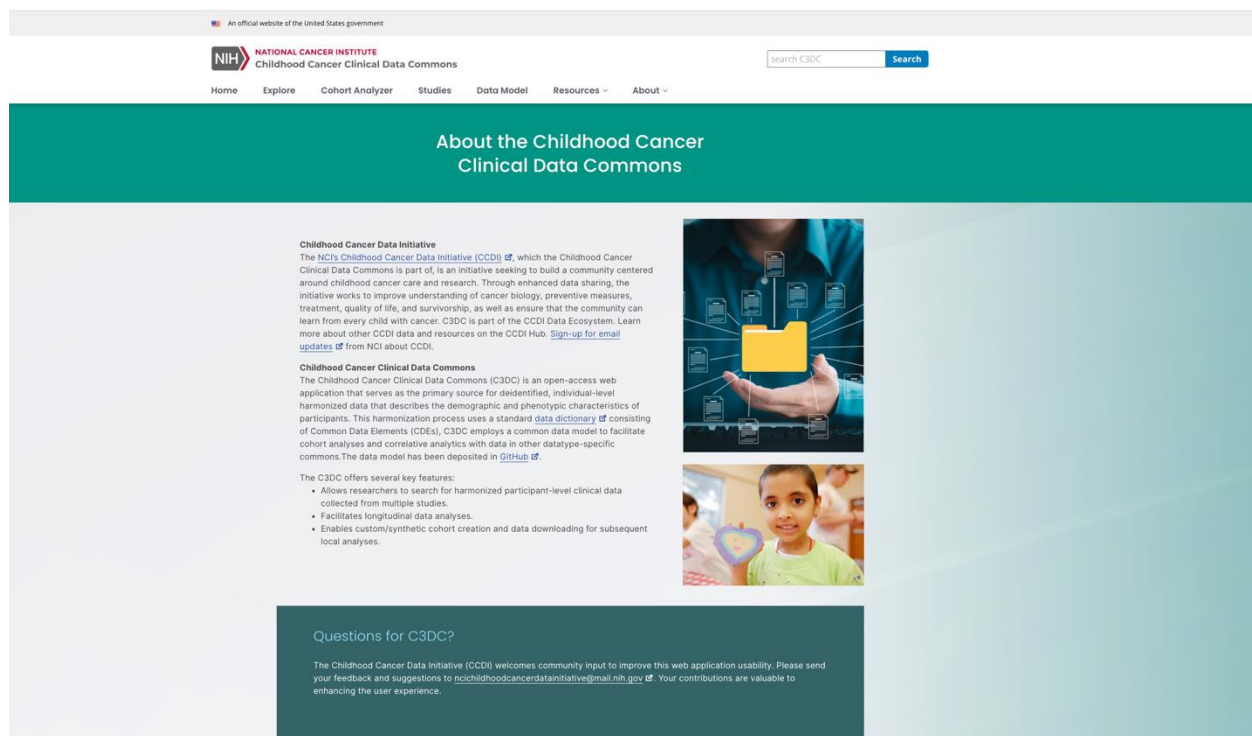
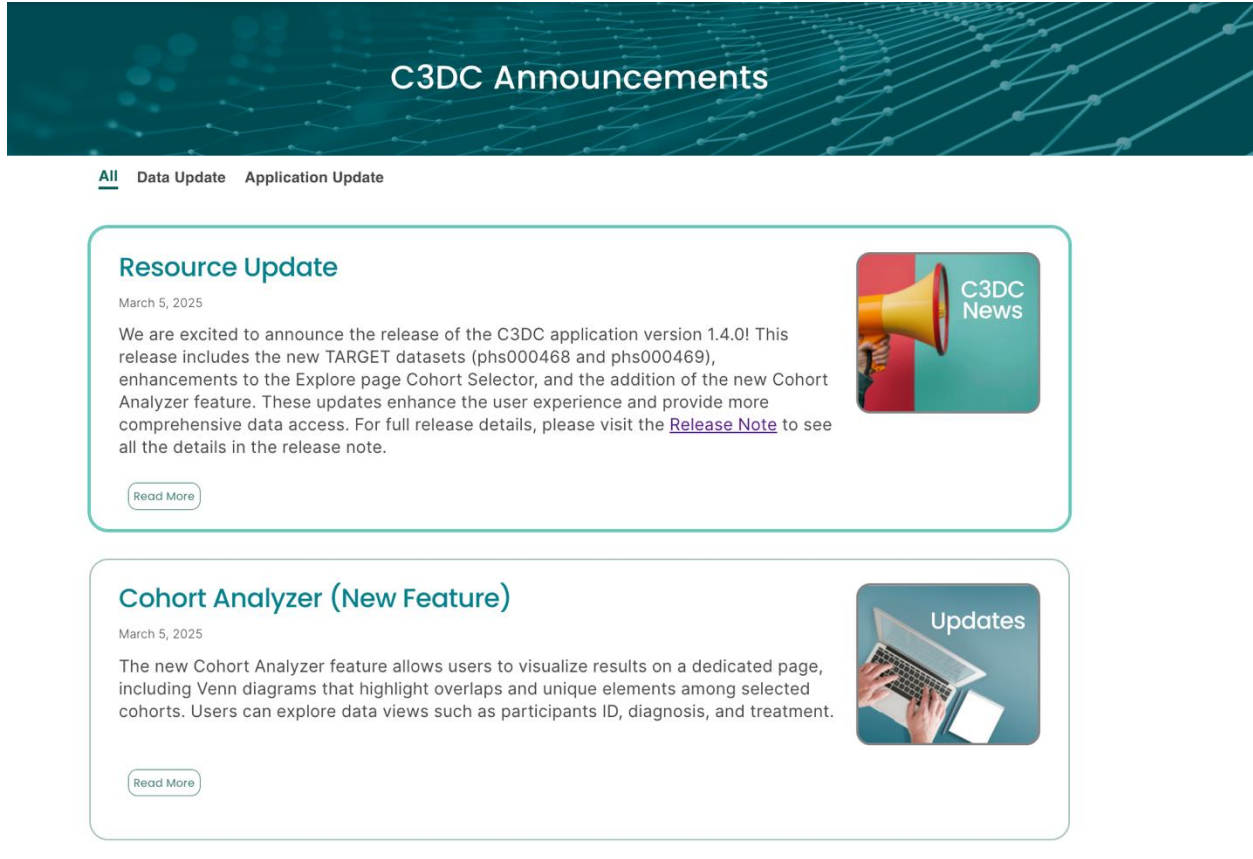


Figure 21: The About page

Announcements

The Announcements page contains all C3DC updates, both data and application, with the newest release information at the top of the list.



The screenshot shows the C3DC Announcements page. At the top is a dark teal header with the text "C3DC Announcements" in white. Below the header is a navigation bar with three links: "All", "Data Update", and "Application Update". The "All" link is underlined. The main content area features two announcement cards. The first card is titled "Resource Update" and dated "March 5, 2025". It contains text about the release of C3DC application version 1.4.0, mentioning new TARGET datasets and enhancements to the Cohort Selector and Cohort Analyzer. A "Read More" button is at the bottom left of the card. To the right of the text is an image of a megaphone with the text "C3DC News". The second card is titled "Cohort Analyzer (New Feature)" and also dated "March 5, 2025". It describes the new Cohort Analyzer feature, which allows users to visualize results on a dedicated page, including Venn diagrams. A "Read More" button is at the bottom left of the card. To the right of the text is an image of a laptop with the text "Updates".

C3DC Announcements

[All](#) [Data Update](#) [Application Update](#)

Resource Update

March 5, 2025

We are excited to announce the release of the C3DC application version 1.4.0! This release includes the new TARGET datasets (phs000468 and phs000469), enhancements to the Explore page Cohort Selector, and the addition of the new Cohort Analyzer feature. These updates enhance the user experience and provide more comprehensive data access. For full release details, please visit the [Release Note](#) to see all the details in the release note.

[Read More](#)

Cohort Analyzer (New Feature)

March 5, 2025

The new Cohort Analyzer feature allows users to visualize results on a dedicated page, including Venn diagrams that highlight overlaps and unique elements among selected cohorts. Users can explore data views such as participants ID, diagnosis, and treatment.

[Read More](#)

Figure 22: The Announcements page

Release Notes

The Release Notes page contains the change logs of all releases, noting the differences between the previous versions of both the data and application.

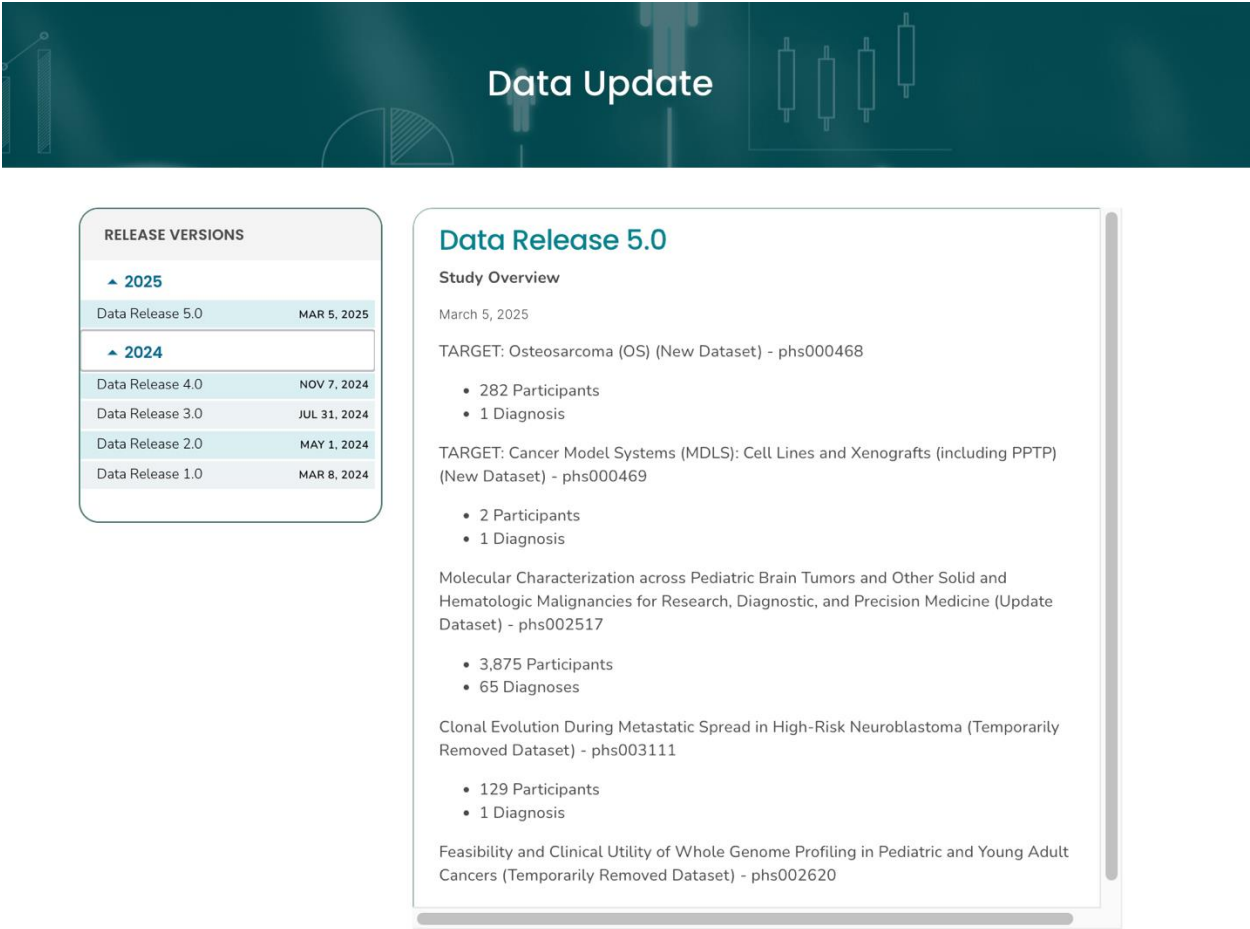


Figure 23: The Release Notes page

User Guide

The User Guide page contains the mission statements for CCDI and C3DC, as well as useful resources related to the C3DC project.

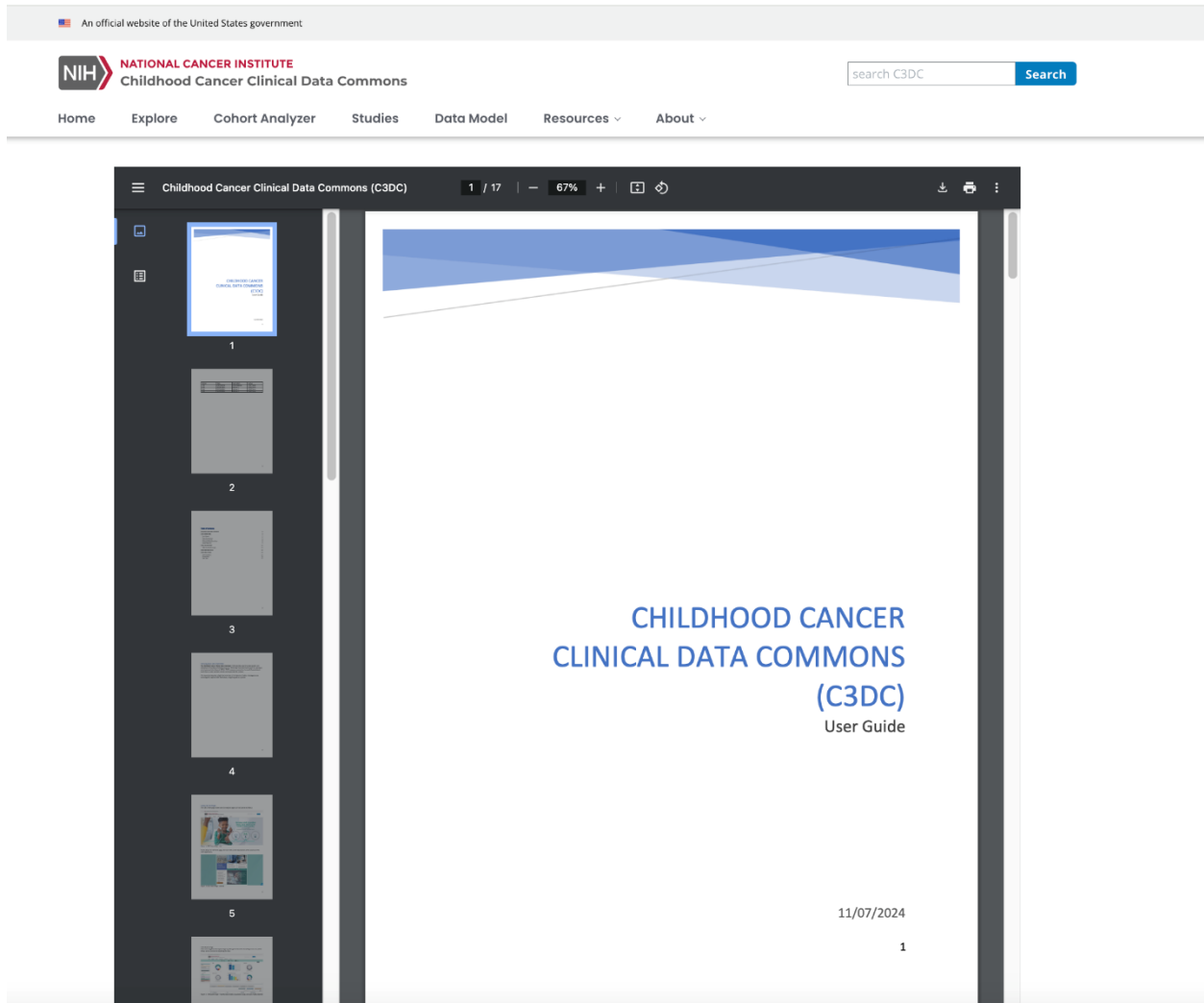


Figure 24: The User Guide