



CHILDHOOD CANCER CLINICAL DATA COMMONS (C3DC)

User Guide

02/4/2026

Version	Date	Description	Author
1.0.0	03/05/2024	Initial Release	C3DC Team
1.1.0	05/01/2024	Release 2	C3DC Team
1.2.0	07/31/2024	Release 3	C3DC Team
1.3.0	11/07/2024	Release 4	C3DC Team
1.4.0	03/05/2025	Release 5	C3DC Team
1.6.0	06/18/2025	Release 6	C3DC Team
1.7.0	10/22/2025	Release 7	C3DC Team
1.8.0	02/04/2026	Release 8	C3DC Team

Table of Contents

INTRODUCTION AND OVERVIEW	4
C3DC HOME PAGE	5
C3DC EXPLORE PAGE	7
C3DC USER GUIDE	8
FACET SEARCH	8
CHART VISUALIZATION	9
TABLE VISUALIZATION	10
SYNONYM SEARCH	11
CPI PARTICIPANT INDICATOR	12
DOWNLOAD HARMONIZED DATA	13
COHORT SELECTOR	13
C3DC COHORT ANALYZER	17
CUSTOMIZABLE PROPERTIES	17
ENHANCED ANALYTICAL CAPABILITIES	18
COHORT ANALYZER TUTORIAL	18
C3DC STUDIES PAGE	23
C3DC STUDIES DETAILS PAGE	23
C3DC DATA MODEL PAGE	25
C3DC RESOURCE PAGE	26
C3DC ABOUT PAGE	27
ANNOUNCEMENTS	29
RELEASE NOTES	30
USER GUIDE	31

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.

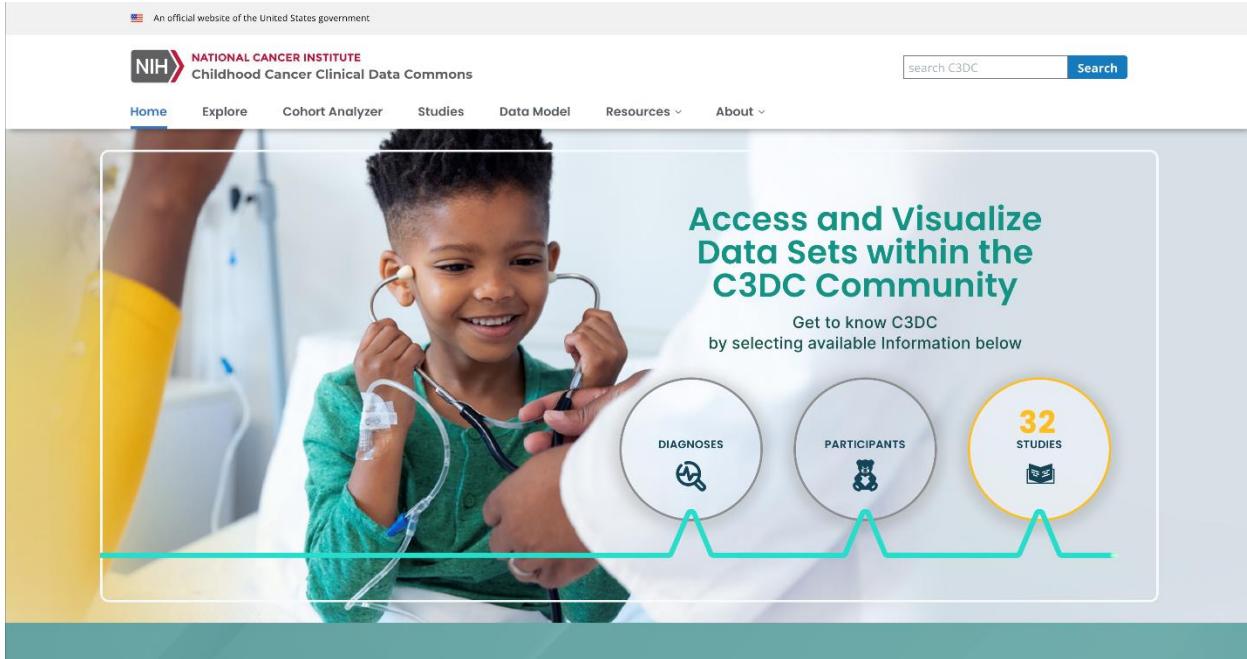


Figure 1: C3DC Home Page (top)

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

The screenshot shows the bottom portion of the C3DC Home Page. It features a grid of four main sections: 'About the Childhood Cancer Clinical Data Commons' (with a network diagram icon), 'Data Model' (with a person using a laptop icon), 'Announcements' (with a person using a laptop icon), and 'C3DC Data' (with a hand interacting with a digital interface icon). Below this grid is a dark blue footer section containing links for 'About', 'Resources', 'Policies', and a 'Sign up for email updates' form.

About the Childhood Cancer Clinical Data Commons

The Childhood Cancer Clinical Data Commons (C3DC) is a database that houses childhood cancer demographics and phenotypic clinical data. 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.

[READ MORE](#)

Data Model

Review components to gain a deeper understanding of the specifics of harmonization.

[EXPLORE DATA MODEL](#)

Announcements

The latest updates for both datasets and the application are now available. Stay current with the latest enhancements.

[EXPLORE ANNOUNCEMENTS](#)

C3DC Data

Explore and export C3DC data for analysis alongside other data types.

[EXPLORE DATA](#)

About

- [About C3DC](#)
- [About CCDI](#)
- [Contact Us](#)
- [Release Notes](#)

Resources

- [C3DC Data Model](#)
- [Bento](#)

Policies

- [Accessibility](#)
- [FOIA](#)
- [Privacy & Security](#)
- [Disclaimer](#)
- [Vulnerability Disclosure](#)

Sign up for email updates

Enter your email address

[Sign up](#)

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.

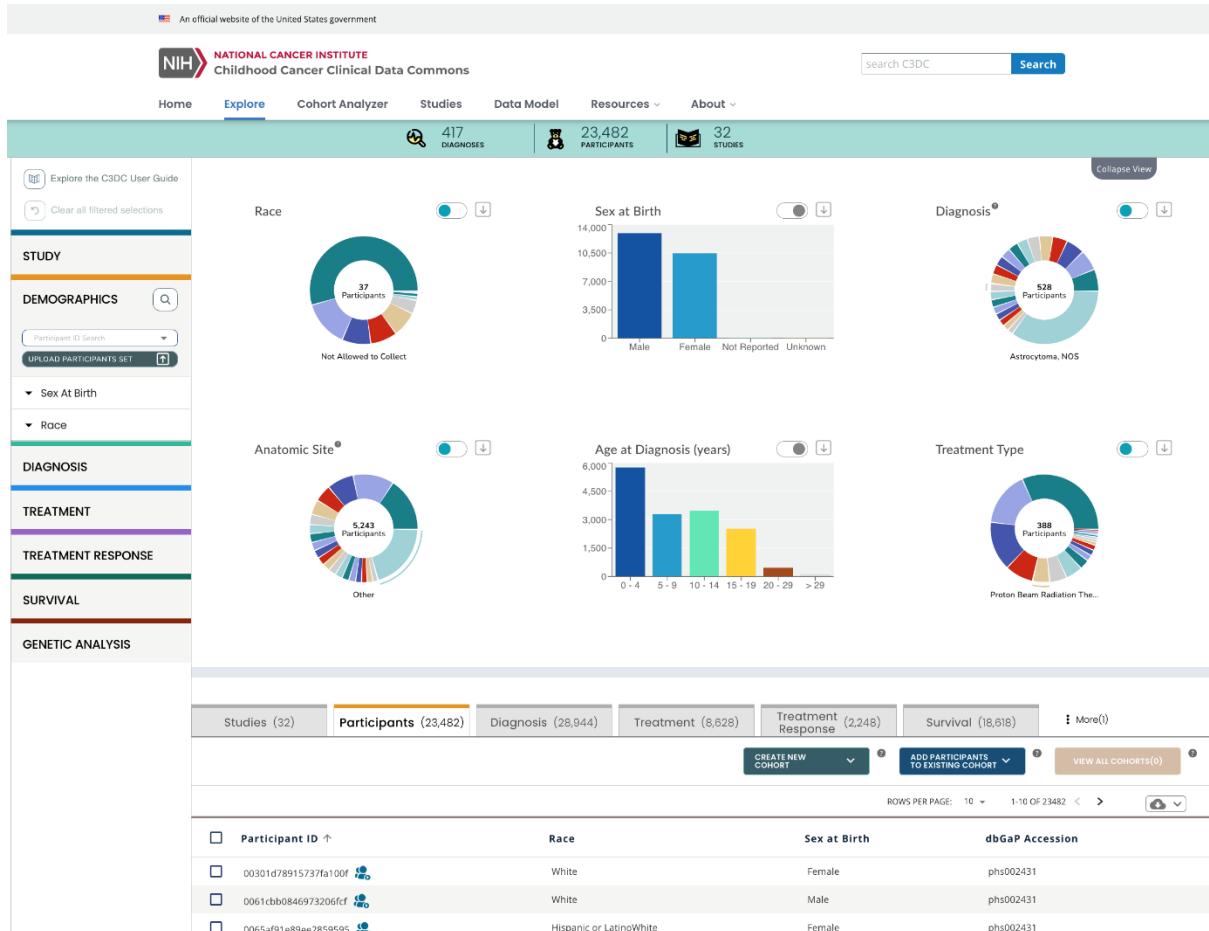


Figure 3: The Explore Page – Faceted Search (left), Visualization (top), and Data Tables (bottom)

C3DC User Guide

The “Explore the C3DC User Guide” button allows a quick access to this C3DC user manual, offering detailed feature explanations and step-by-step instructions. Additionally, users will find various use cases in this guide, making it easy to find help and learn how to navigate the system effectively.

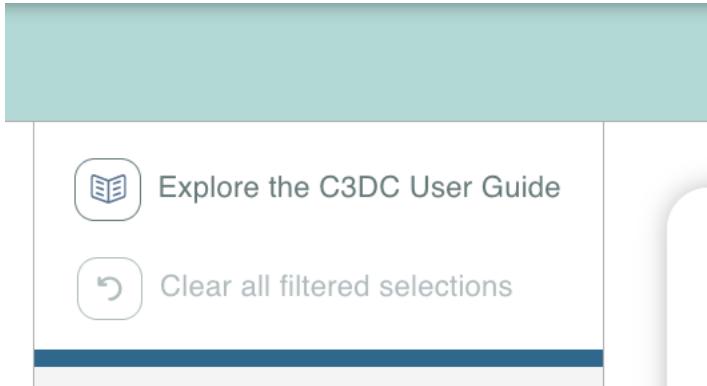


Figure 4: C3DC User Guide Button in the Explore Page

Facet Search

Users can select a subset of the childhood cancer participants by choosing filtering options from drop-down lists within seven main categories: Study, Demographics, Diagnosis, Treatment, Treatment Response, Survival, and Genetic Analysis. Under each main category, there are multiple subcategories available for users to perform refined filtering.

Figure 5: 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 6: Results returned from a faceted search in the Stats bar

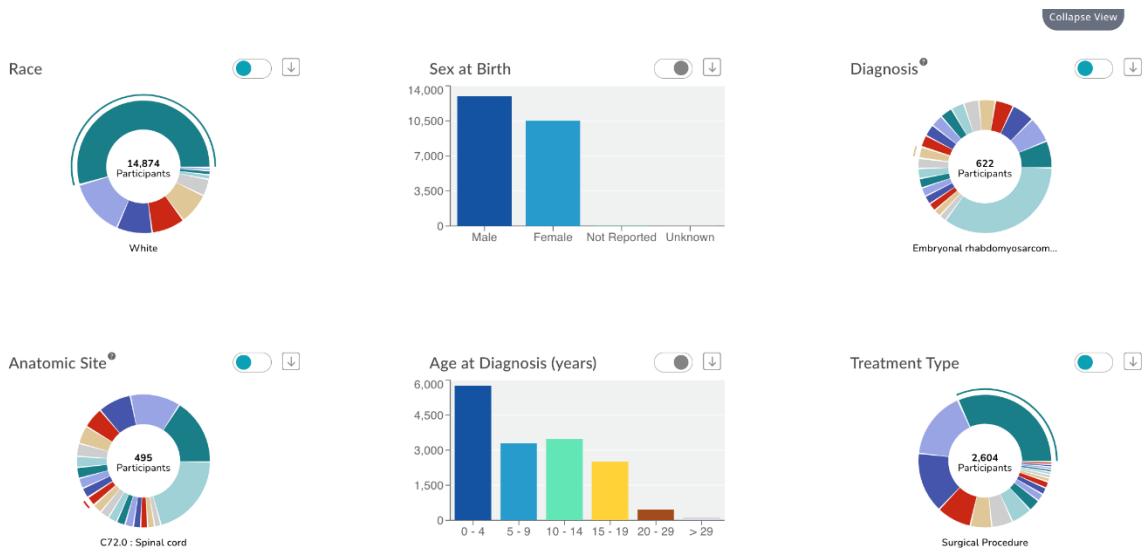


Figure 7: Results returned from a faceted search in six graphs in the Visualization section

- Chart visualization illustrates the proportions of each value in Race, Sex at Birth, Diagnosis, Anatomic Site, Age at Diagnosis, and Treatment Type in the form of either pie chart or histogram.
- Users can easily switch between pie chart and histogram through the toggle button by each visualization.
- Download button for each visualization is also available for users to download a PNG file.

Table Visualization

Table visualization displays participant details with **tooltips explaining table headers**, including Studies, Participants, Diagnosis, Treatment, Treatment response, Survival, and Genetic Analysis. Users can click “More” button on the upper right side to access Genetic Analysis table.

Count of Participant Record

Participant ID ↑	Race	Sex at Birth	dbGaP Accession
00301d78915737fa100f	White	Female	phs002431
0061ccb0846973206fcf	White	Male	phs002431
0065af91e89ee2859595	Hispanic or LatinoWhite	Female	phs002431
008b04a84717e7d007a4	White	Male	phs002431
009f9efadc9602a28fb	White	Female	phs002431
00c5e4372375eb3627f3	White	Female	phs002431
00dfd38d509984eda22	White	Male	phs002431
00e3d1d383c8c08a25d2	White	Male	phs002431
00eff50ac98a200da9f9	White	Male	phs002431
01504e3b7031a9eec2b2	White	Female	phs002431

Figure 8A: Results returned from a faceted search in the Table section

Genetic Analysis (90,123)

Participant ID ↑	Genetic Analysis ID	Gene Symbol	Status	Reported Significance	Reported Significance System	HGVS Genome	Alteration	dbGaP Accession
PANLMU	PANLMU_0DHY47_TumorNormal_v1_genetic_analysis_764313	APOBEC3B, BCR, CHEK2, CLTC1, DGC88, CCNC, EPHA7, FOXO3, GOPC, HACE1,...	Present	Tier 1A	ACMG/GCG	Not Reported	Whole Arm Loss	phs002790
PANLMU	PANLMU_0DHY47_TumorNormal_v1_genetic_analysis_586416		Present	Tier 1A	ACMG/GCG	Not Reported	Loss	phs002790

Figure 8B: Genetic analysis table in the Table section

Synonym Search

When searching for a participant, Users can use any recognized synonym from the CCDI Participant Index (CPI). The system will identify the synonym, note that it matches a known participant, and display the corresponding participant_id and details in the results table below.

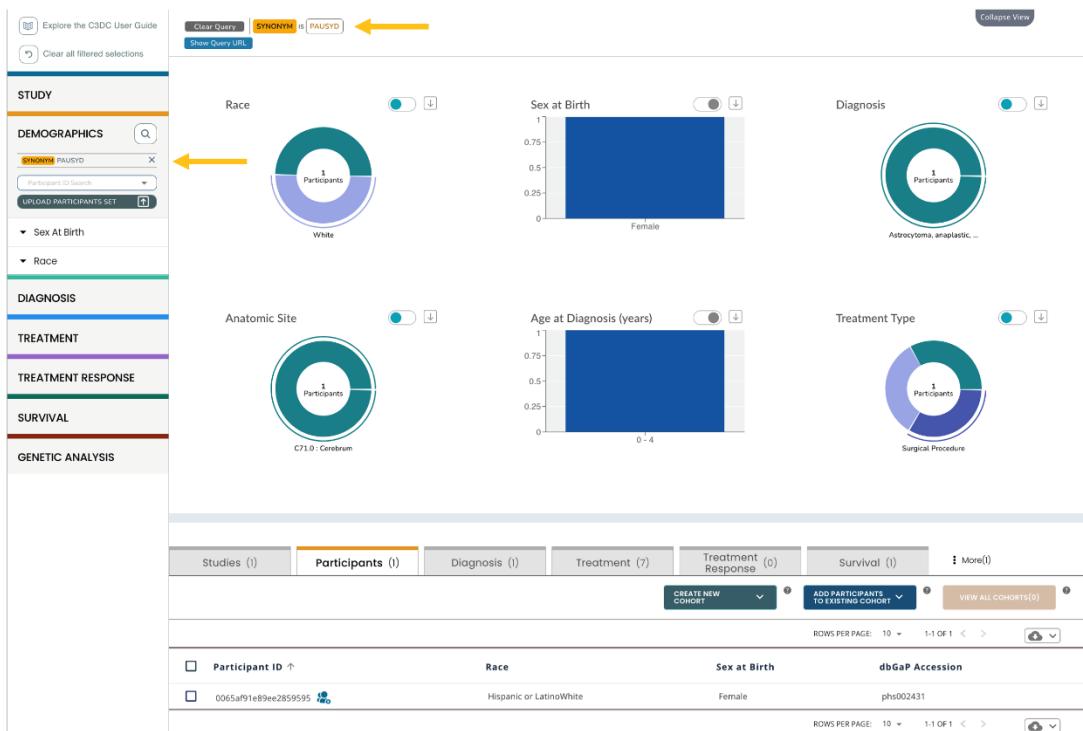


Figure 9: Searching by synonym capabilities and match confirmation

CPI Participant Indicator

For participants with additional synonym values from the CPI, an icon will be displayed next to the participant_id values within the Participant tab table. Hovering over the icon displays a tooltip with the clickable link that opens a popup table containing synonym information from the CPI.

Studies (32)	Participants (23,482)	Diagnosis (28,944)	Treatment (8,628)	Treatment Response (2,248)	Survival (18,618)	More(1)
CREATE NEW COHORT ADD PARTICIPANTS TO EXISTING COHORT VIEW ALL COHORTS(0)						
ROWS PER PAGE: 10 1-10 OF 23482 < >						
<input type="checkbox"/> Participant ID ↑	Race	Sex at Birth	dbGaP Accession			
<input type="checkbox"/> 00301d78915737fa100f	White	Female	phs002431			
<input type="checkbox"/> 0065af91e89ee2859595	White	Male	phs002431	All mapped identifiers in the CCDI Participant Index (CPI) are available here .		
<input type="checkbox"/> 0065af91e89ee2859595	Hispanic or LatinoWhite	Female	phs002431			
<input type="checkbox"/> 008b04a84717e7d007a4	White	Male	phs002431			

Figure 10A: Table view with annotations showing the existence of synonyms for a participant.

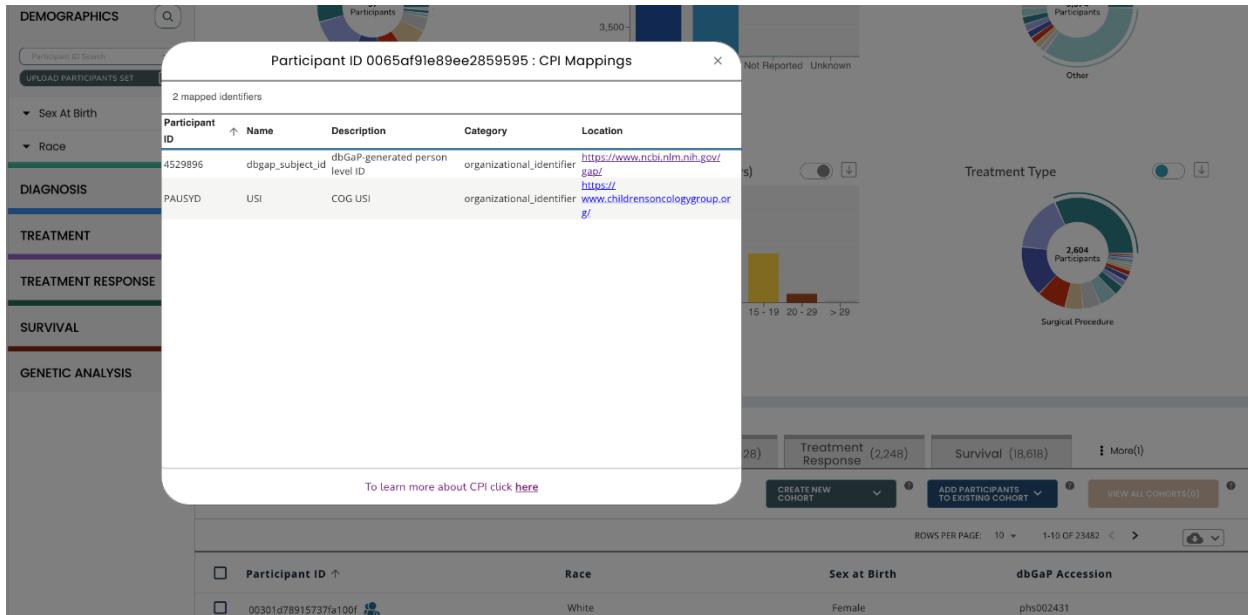


Figure 10B: Clicking on participant icon shows a pop-up window showing synonym mapping

Download Harmonized Data

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



Figure 11A: Download button to download harmonized data in the Explore Page

A1	Participant Id	Diagnosis ID	Diagnosis	Diagnosis Cl	Diagnosis Ba	Diagnosis Cc	Disease	Phenotype	Tumor Class	Anatomic Site	Age at Diagnx	Toronto	Child	Tumor Grade	Tumor Stage	Tumor Stage	Tumor Stage	Study ID	Dbgap Accession
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R		
1	Participant Id	02296835-e61b-4b87-a9fa-fd105b3977bc	042111:Plorc	ICD-O-3.2	Clinical	Not Reported	Initial	Diagn Primary	C71.6: Cerel	2005	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002790.v phs002790	
2	PBCHDL	0742008ec41e722cad36558be43221_1	0742008ec41e722cad36558be43221_1	High-Grade C NS	Diagn Pathological	High-grade gli	Initial	Diagn Primary	C71.6: Cerel	272	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517 phs002517	
3	PT_17H2AAS	0beef93aae1635b5952ba05362280cb5d_1	040000: Astro	ICD-O-3.2	Pathological	Low-grade gli	Initial	Diagn Primary	C71.6: Cerel	1548	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517	
4	PT_RC1H1QJ	0e958c55474d5e655083444fcf825e7_1	0e958c55474d5e655083444fcf825e7_1	Medulloblast	CNS Diagn Pathological	Medulloblast	Initial	Diagn Primary	C71.6: Cerel	2212	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517 phs002517	
5	PT_V9WP047R	0f04db0e30132b27c7a7669887879026_5	0f04db0e30132b27c7a7669887879026_5	094000: Astro	ICD-O-3.2	Pathological	Low-grade gli	Initial	Diagn Primary	C71.6: Cerel	1805	Not Reported	phs002517 phs002517						
6	PT_D2867YXN6	0f0e69d0f3e7b9c03dac0b6f9ca87513_1	0f0e69d0f3e7b9c03dac0b6f9ca87513_1	095083: Atypic	ICD-O-3.2	Pathological	Atypical Tera	Post-Mortem	Primary	C71.6: Cerel	3562	Not Reported	phs002517 phs002517						
7	PT_6NB2556	0f0e640d84c553f8ee48456825c57a_1	0f0e640d84c553f8ee48456825c57a_1	Low-Grade G1	CNS Diagn Pathological	Low-grade gli	Progression	Primary	C71.6: Cerel	2402	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517 phs002517	
8	PT_962TCBVR	1466dd5dc5a7fb5e93717270f030db_1	1466dd5dc5a7fb5e93717270f030db_1	Low-Grade G1	CNS Diagn Pathological	Low-grade gli	Initial	Diagn Primary	C71.6: Cerel	1372	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	Not Reported	phs002517 phs002517	
9	PT_GAKHIC0Y2	1625090f2e143775606a53deaa3734_1	1625090f2e143775606a53deaa3734_1	940000: Astro	ICD-O-3.2	Pathological	Low-grade gli	Initial	Diagn Primary	C71.6: Cerel	3698	Not Reported	phs002517 phs002517						
10	PT_6VNN63H	1801de1a7eb6342681_1	1801de1a7eb6342681_1	143775606a53deaa3734_1	143775606a53deaa3734_1	940000: Astro	ICD-O-3.2	Pathological	Low-grade gli	Initial	Diagn Primary	C71.6: Cerel	1473	Not Reported	phs002517 phs002517				
11	PT_64JHC2BH	1b5550275cda4089fd68570602e2b4_1	1b5550275cda4089fd68570602e2b4_1	185550275cda4089fd68570602e2b4_1	185550275cda4089fd68570602e2b4_1	940000: Astro	ICD-O-3.2	Pathological	Low-grade gli	Initial	Diagn Primary	C71.6: Cerel	2326	Not Reported	phs002517 phs002517				
12	PT_KJKY2JTZ	1b55db2568dd37deb896fa1a6c9ac25_1	1b55db2568dd37deb896fa1a6c9ac25_1	1b55db2568dd37deb896fa1a6c9ac25_1	1b55db2568dd37deb896fa1a6c9ac25_1	Low-Grade G1	CNS Diagn Pathological	Low-grade gli	Recurrent	Di Primary	C71.6: Cerel	1717	Not Reported	phs002517 phs002517					
13	PT_962TCBVR	19c:b19-fac-4661-a060-0b9274d00d5	19c:b19-fac-4661-a060-0b9274d00d5	9421:1:Plorc	ICD-O-3.2	Clinical	Not Reported	Initial	Diagn Primary	C71.6: Cerel	1040	Not Reported	phs002790.v phs002790						

Figure 11B: Download Diagnosis Table tab as CSV file format

Cohort Selector

The Cohort Selector enables users to create a cohort with a size of up to 4000 and manage up to 20 cohorts. This feature offers flexibility to researchers, allowing them to create cohort groups according to their specific requirements.



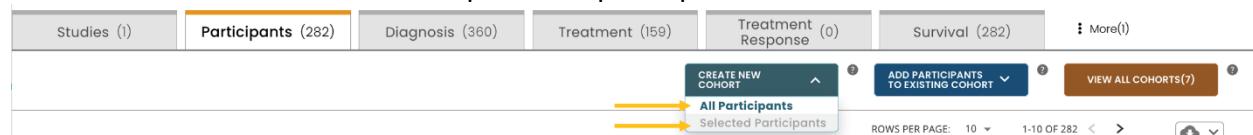
Studies (2)	Participants (1,239)	Diagnosis (1,239)	Treatment (1,186)	Treatment Response (0)	Survival (1,239)	More(?)
				<div style="border: 1px solid orange; padding: 2px;"> CREATE NEW COHORT ADD PARTICIPANTS TO EXISTING COHORT VIEW ALL COHORTS(0) </div>		

Participant ID	Treatment ID	Age at Treatment Start	Age at Treatment End	Treatment Type	Treatment Agent	dbGaP Accession
R10130226	153ed358-b72a-54f5-af2e-f606838917dc	Not Reported	Not Reported	Unknown	Not Reported	phs002518
R10332595	cf89b53c-3bf8-5e70-980d-21d5fdb3233	Not Reported	Not Reported	Not Reported	Not Reported	phs002518
R10400226	92b142b2-8bc0-57e1-8ab6-7dbae30c2dcc	Not Reported	Not Reported	Chemotherapy	Not Reported	phs002518
R10529639	73b8284c-4246-5575-ad05-8742c21b0b09	Not Reported	Not Reported	Surgical Procedure	Not Reported	phs002518
R10572154	16e5705f-4e78-5370-8078-0cb82f6a9e76	Not Reported	Not Reported	Not Reported	Not Reported	phs002518
R10604930	dc74baa2-6d16-5bc9-9d7d-cf9423b1bb42	Not Reported	Not Reported	Unknown	Not Reported	phs002518
R10610670	eac554d1-57f1-5409-b0ab-fda12bdf44f0	Not Reported	Not Reported	Chemotherapy	Not Reported	phs002518

Figure 12A: Cohort Selection features visible on the Explore page Table tabs (see choice buttons in orange rectangle)

Users can do the following:

- Create New Cohort:
 - Users can select participant IDs from the table or choose to add all participants based on the faceted results and create a new cohort.
 - Users can name and describe the cohort for easy reference.
 - A user can add up to 4000 participants in each cohort.



Studies (1)	Participants (282)	Diagnosis (360)	Treatment (159)	Treatment Response (0)	Survival (282)	More(?)
	<div style="border: 1px solid orange; padding: 2px;"> CREATE NEW COHORT All Participants Selected Participants </div>				<div style="border: 1px solid orange; padding: 2px;"> ADD PARTICIPANTS TO EXISTING COHORT VIEW ALL COHORTS(?) </div>	

Participant ID	Race	Sex at Birth	dbGaP Accession
TARGET-40-0A4HLD	Not Reported	Male	phs000468
TARGET-40-0A4HMC	White	Male	phs000468
TARGET-40-0A4HXB	White	Male	phs000468
TARGET-40-0A4HXO	Unknown	Male	phs000468

Figure 12B: Create a new cohort by adding all the participants in the table or selecting a subset of participants using checkbox

- Add Participants to Existing Cohort:
 - Select participants and add them to existing cohorts or remove them. Entire cohorts can also be deleted as needed.

The screenshot shows a participant management interface with several tabs at the top: Studies (1), Participants (403) [highlighted in orange], Diagnosis (403), Treatment (0), Treatment Response (0), Survival (403), and More(1). Below the tabs is a search bar and a 'CREATE NEW COHORT' button. A dropdown menu titled 'ADD PARTICIPANTS TO EXISTING COHORT' is open, showing 'All Participants' and 'Selected Participants'. Under 'Selected Participants', 'phs001437' is checked. To the right of the dropdown is a 'VIEW ALL COHORTS(1)' button. A yellow arrow points from the text 'Selected Participants' in the dropdown to the checked checkbox next to 'phs001437'. Below the dropdown is a table with columns: Participant ID, Race, Sex at Birth, and dbGaP Accession. The table lists six participants, with PAJALK and PAJALV having their checkboxes checked.

Participant ID	Race	Sex at Birth	dbGaP Accession
PAIZZZ	White	Female	phs000720
PAJAGJ	White	Female	phs000720
PAJAKA	White	Female	phs000720
<input checked="" type="checkbox"/> PAJALK	Black or African American	Female	phs000720
<input checked="" type="checkbox"/> PAJALV	White	Male	phs000720

Figure 12C: Adding entire or selected participants to an existing cohort

- View All Cohort(s): View a list of all created cohorts, making it easier to manage and analyze groups.
 - Cohort ID: Create your own IDs to identify saved cohorts
 - Cohort Description: Create descriptions for saved cohorts
 - Save Changes: Save the changes made to the selected cohort. This includes changes to cohort ID, cohort description, and any participants.
 - Download Selected Cohort:
 - Download the metadata of selected cohort in one of two formats.
 - Manifest CSV: a list of participant IDs and high-level metadata.
 - Metadata JSON: a JSON file containing all metadata information for the participants in the selected cohort, including CPI synonyms
 - View Cohort Analyzer: Navigate to the Cohort Analyzer from the cohort list.
 - Explore in CCDI Hub: Export cohorts (up to 4000 participants for each cohort) that open the CCDI Hub with pre-filtered data based on selected participants.

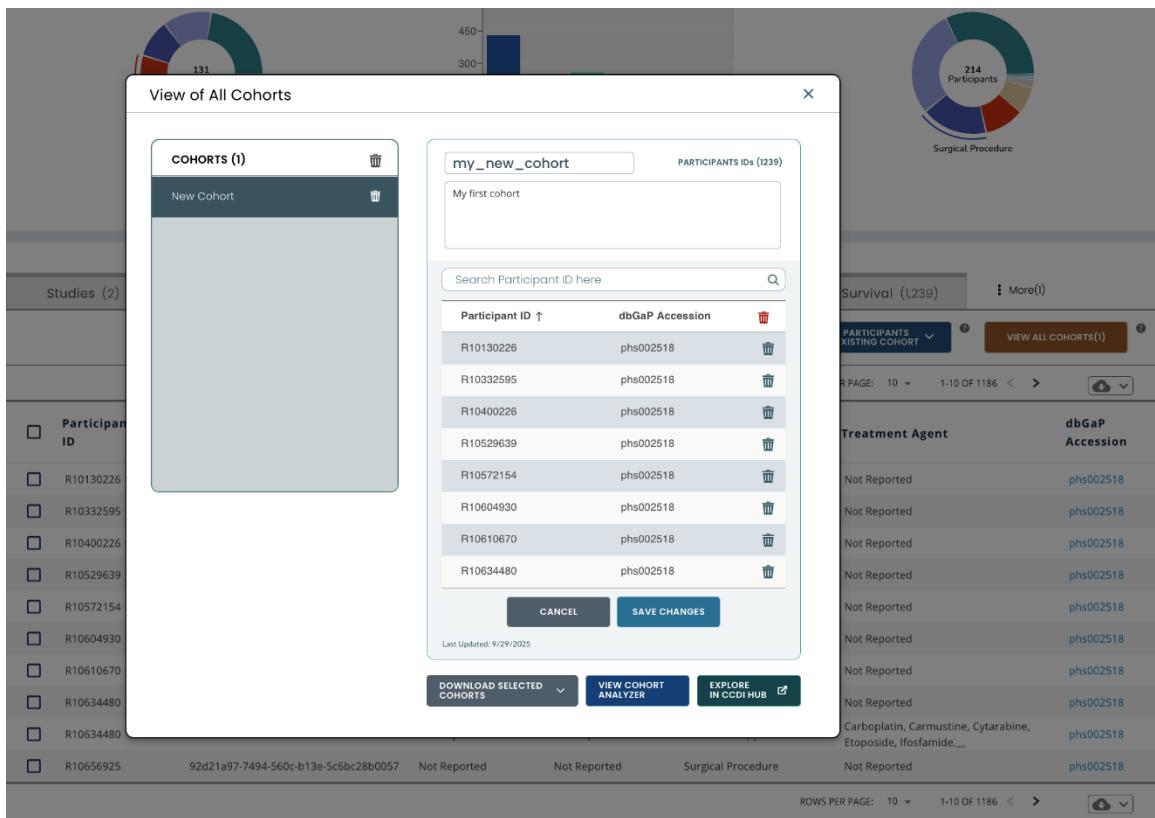


Figure 12D: View All Cohorts popup allows users to manage up to 20 cohorts. Users were given choices to download cohort metadata, view cohort analyzer, and export cohort to CCDI Hub

C3DC Cohort Analyzer

The Cohort Analyzer offers a powerful method to explore how various clinical attributes overlap and differ across multiple groups. The Cohort Analyzer is designed to compare up to three cohorts and visualize their intersections through an interactive Venn diagram, corresponding histograms, and a data table. This feature leverages cohorts created on the Explore page, enabling users to analyze key relationships and distinctions based between datasets effectively.

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.

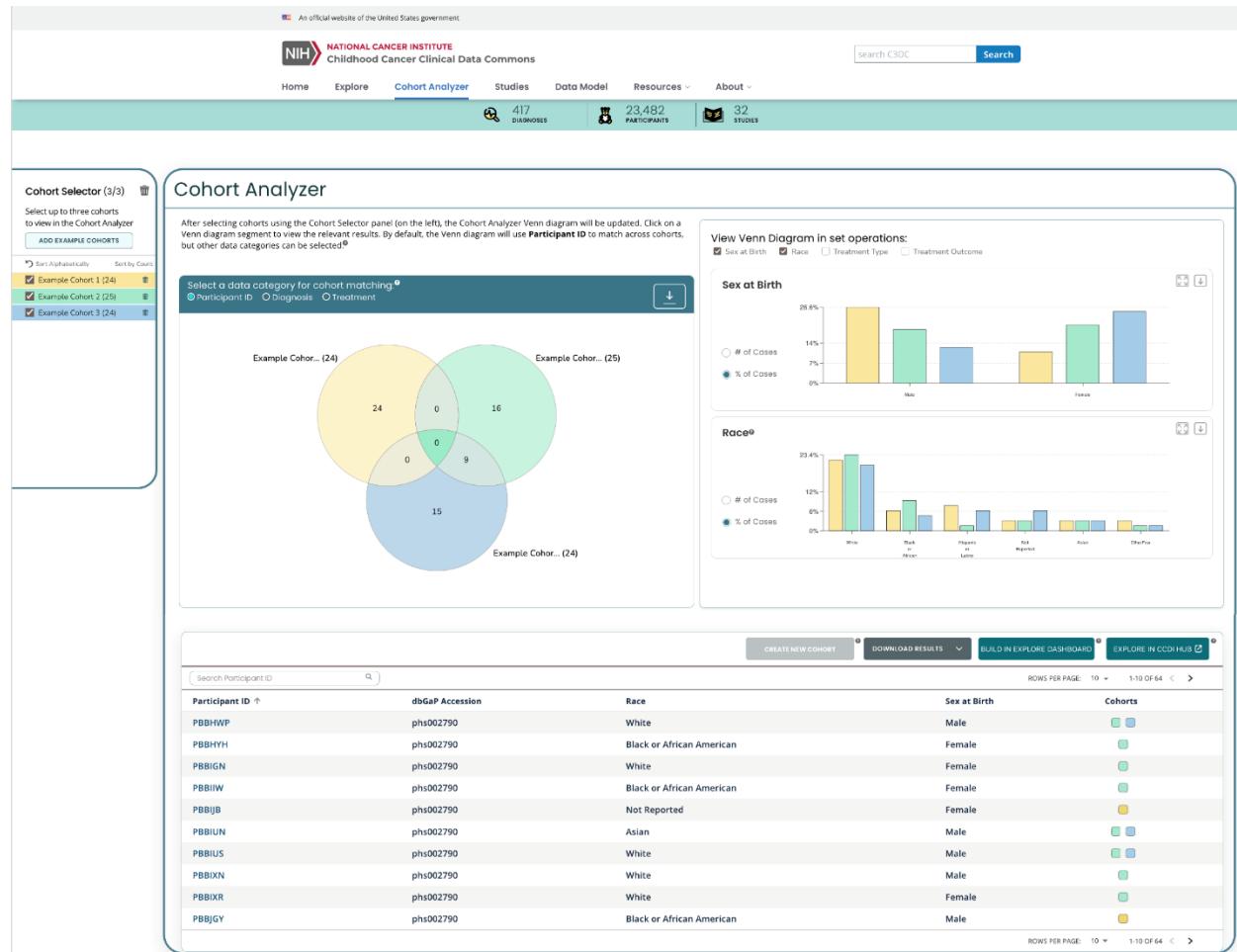


Figure 13: Cohort Analyzer landing page

Customizable Properties

The radio buttons allow users to select more than one property for comparison. The Venn diagram of Participant ID shows the number of participants shared between different sets, while

the Venn diagrams of Diagnosis or Treatment display the number of unique values under each category. 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, histograms, 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.
- Download result: The cohort result can be downloadable as a CSV with individual high-level metadata or a JSON file with comprehensive metadata, including CPI synonyms.
- Build in Explore Dashboard: Export your analysis into a pre-filtered view within the Explore Dashboard for streamlined review and exploration.
- Explore in CCDI Hub: Export cohorts (up to 4000 participants) that open the CCDI Hub with pre-filtered data based on selected participants.
- “Add Example Cohorts” button allows user to explore cohort analyzer features easily by adding 3 mock cohorts

Cohort Analyzer Tutorial

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. Histograms section will automatically display histograms of Sex at Birth, and Race by default. Users can add Treatment Type and Treatment Outcome histograms by checking the checkboxes above the histograms. In this example, we are using the participant ID and diagnosis, thus the table will contain properties specific to the participant as well as show to what cohorts the participant belongs.

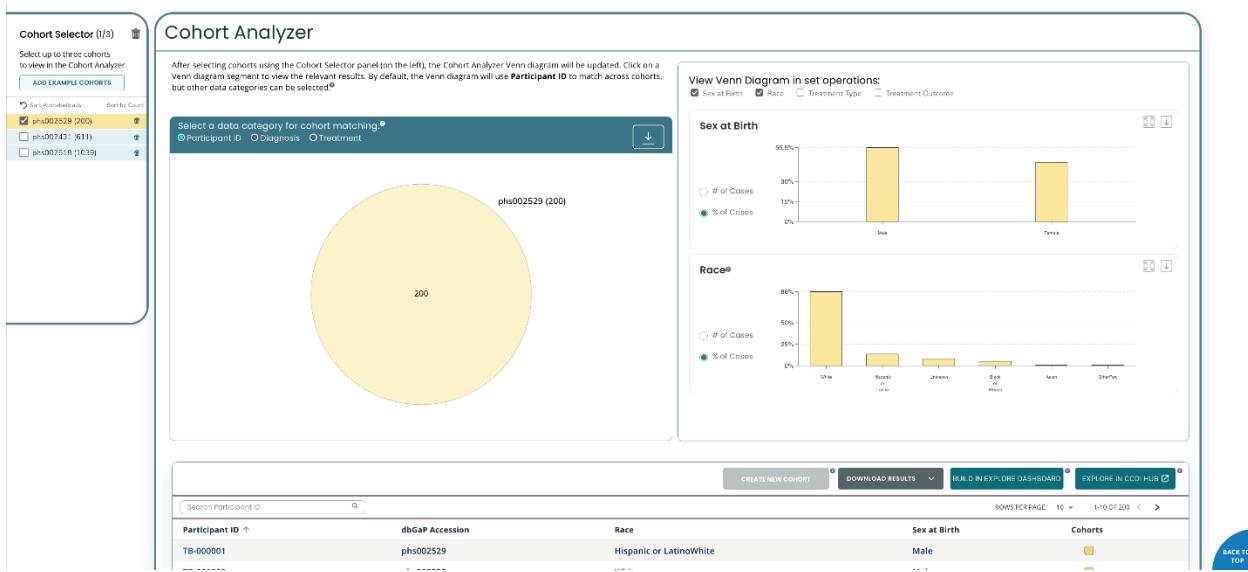


Figure 14: One Cohort Selected: You have created and selected one cohort.

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. Clicking the Diagnosis radio button shows a Venn diagram of unique and shared diagnosis values between two cohorts. In the table below, with none of the Venn diagram selected, it will display all participants and their respective cohort. Selecting part of the Venn diagram will update the table content below accordingly. Histograms on the right side also updates automatically comparing two selected cohorts side-by-side.

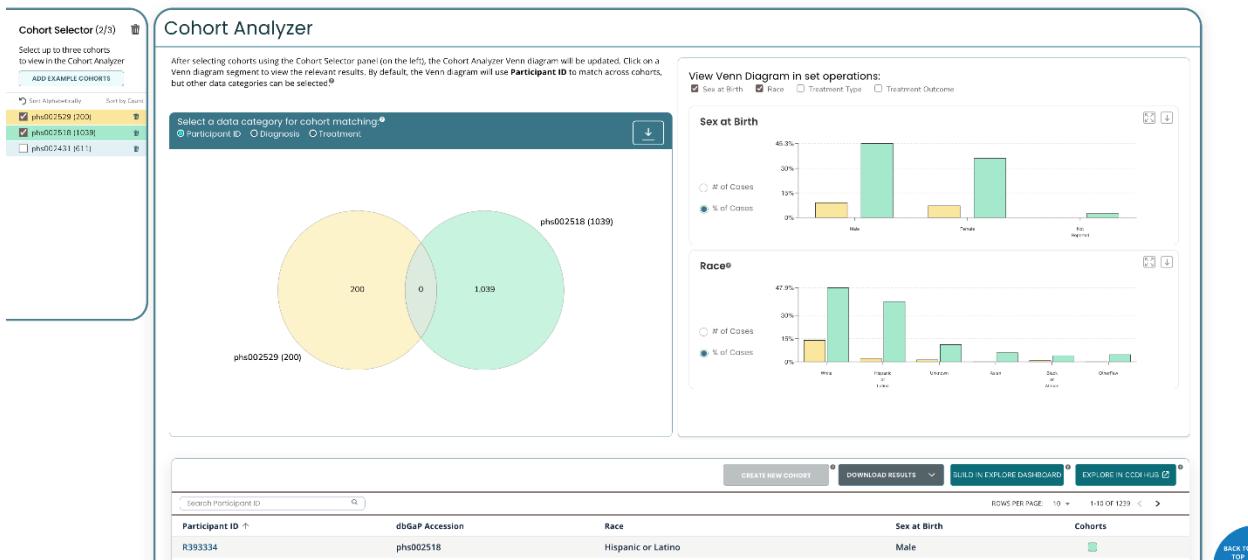


Figure 15: Two Cohorts Selected. You have selected two cohorts. Visualize shared and unique data points between these cohorts.

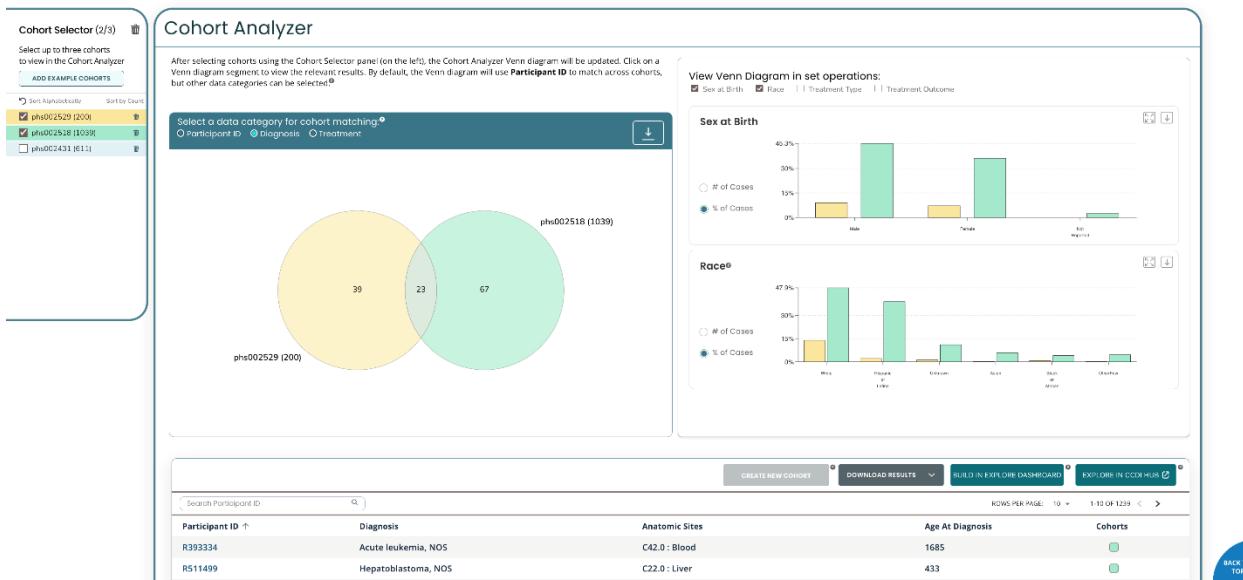


Figure 16: Two Cohorts Selected. Visualize shared and unique Diagnosis values between the two selected cohorts

Select a third and final cohort. The Venn diagram and table will update again to display all participant-level data corresponding to the radio button selection.

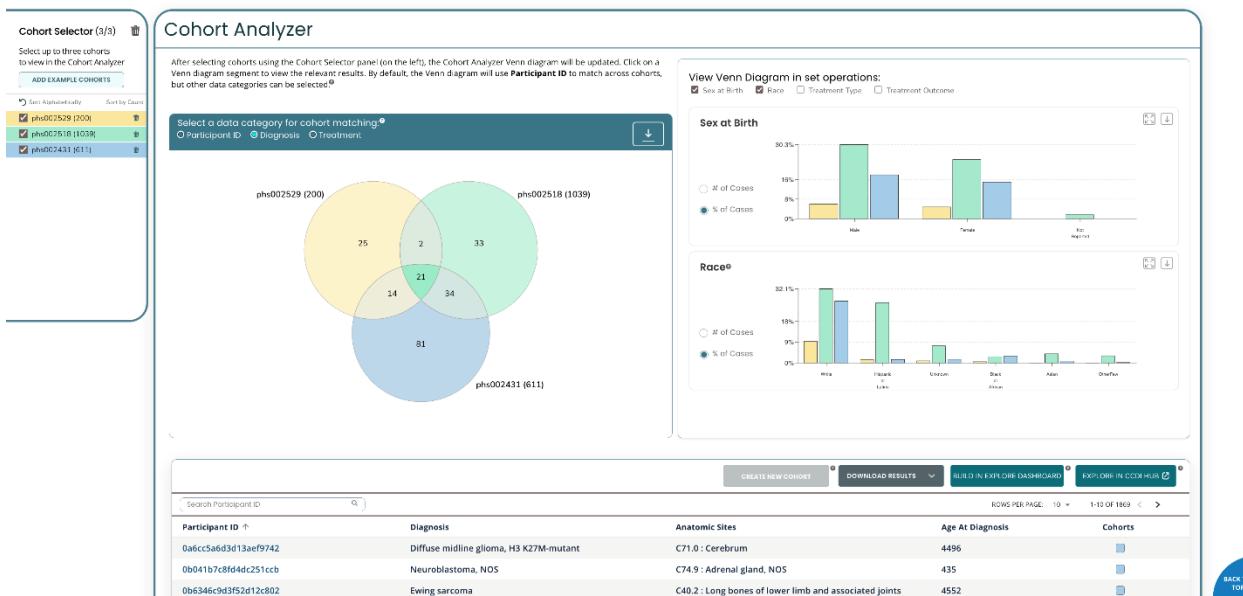


Figure 17: Three Cohorts Selected: You have selected three cohorts. Explore their intersections and unique attributes using the Venn diagram.

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.

At this point, you can select one of these pieces on the Venn diagram to update the table to show only those participants and their respective data. In the example below, the center intersection was selected. The table updates showing only participants that are found in all three cohorts. With this section selected, a user can also create an entirely new cohort with these filtered participants by clicking the “Create New Cohort” button.

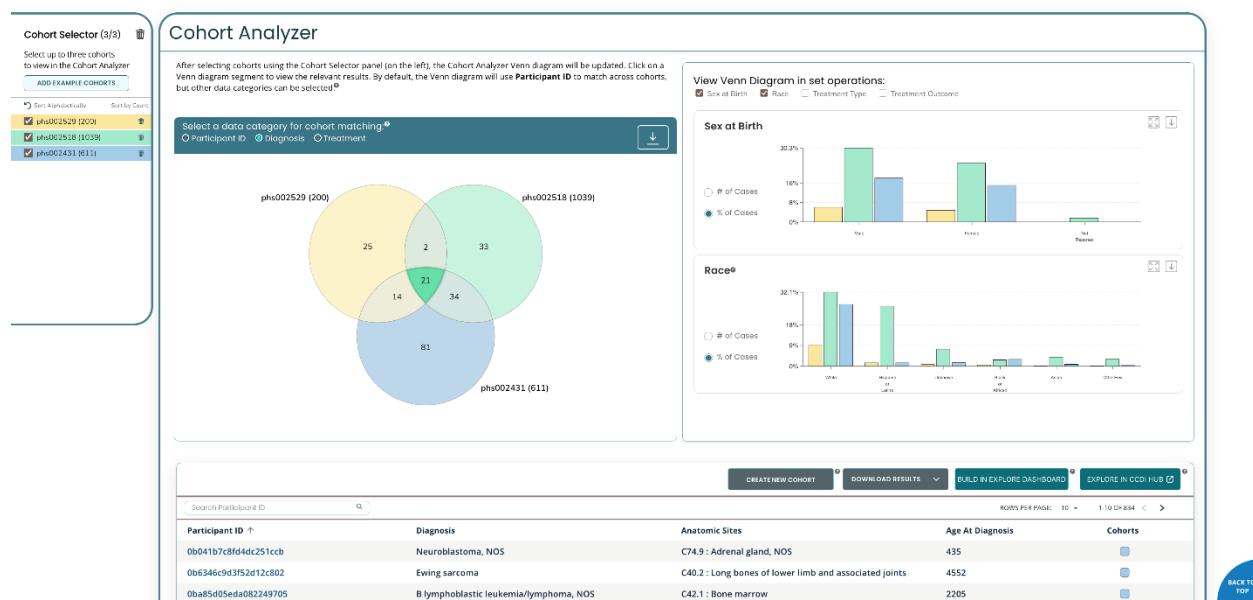


Figure 18: View the center intersection between all Cohort selected (see dark green highlighted region)

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.

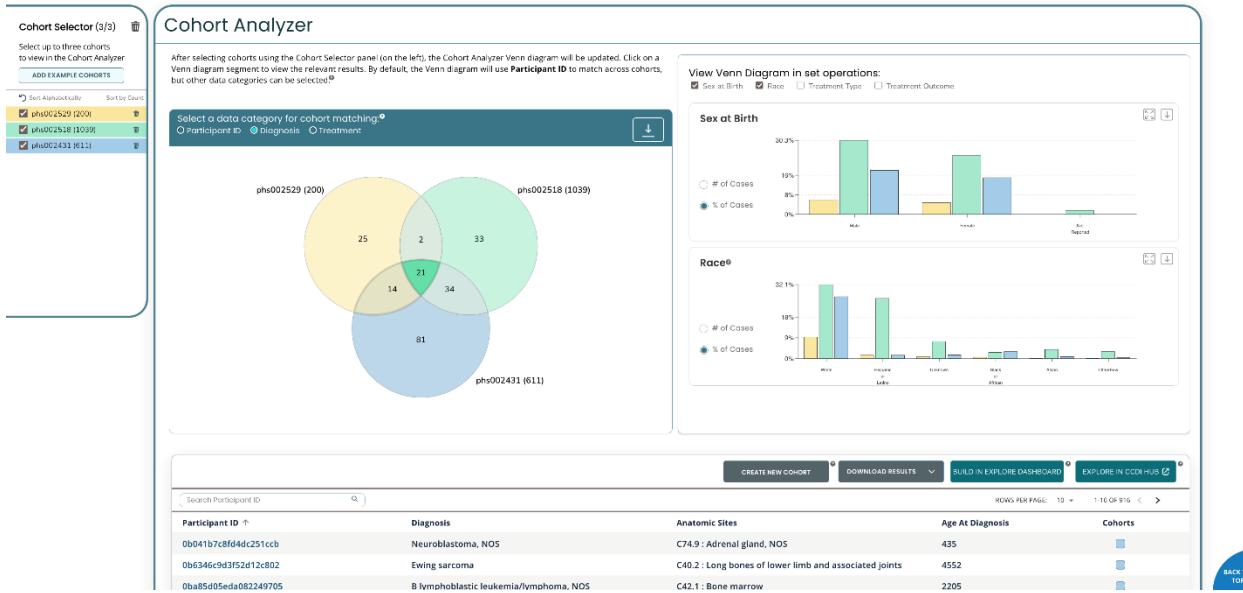


Figure 19: View the specific intersections between selected Cohorts (see dark green and gray highlighted area)

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.

The screenshot shows the C3DC Studies Page. At the top, there is a navigation bar with links for Home, Explore, Cohort Analyzer, Studies (which is the active tab), Data Model, Resources, and About. Below the navigation bar, there are three summary statistics: 417 DIAGNOSES, 23,482 PARTICIPANTS, and 32 STUDIES. The main content area is titled "C3DC Studies" and displays a table of study details. The table has columns for Study Name, Participants Count, Diagnosis Count, and dbGaP Accession. The data is paginated with 10 rows per page, showing results 1-10 of 32. The table includes the following data:

Study Name	Participants Count	Diagnosis Count	dbGaP Accession
TARGET: Acute Lymphoblastic Leukemia (ALL) Pilot Phase 1	231	2	phs000463
TARGET: Acute Lymphoblastic Leukemia (ALL) Expansion Phase 2	1,704	8	phs000464
TARGET: Acute Myeloid Leukemia (AML)	2,144	10	phs000465
TARGET: Kidney, Clear Cell Sarcoma of the Kidney (CCSK)	13	1	phs000466
TARGET: Neuroblastoma (NBL)	1,119	3	phs000467
TARGET: Osteosarcoma (OS)	282	1	phs000468
TARGET: Cancer Model Systems (MDLS): Cell Lines and Xenografts (including PPTP)	2	1	phs000469
TARGET: Kidney, Rhabdoid Tumor (RT)	69	1	phs000470
TARGET: Kidney, Wilms Tumor (WT)	652	1	phs000471
Genomic Sequencing and Spatial Transcriptome Analysis of Pediatric Rhabdomyosarcoma	403	5	phs000720

Figure 20: The Studies Page

C3DC Studies Details Page

By clicking on the dbGaP accession number (e.g., phs000463), 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 phs000463, phs000464, phs000465, 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](#).

The screenshot shows the C3DC Studies Details page for DBGAP ACCESSION: phs000467. At the top, there's a navigation bar with links for Home, Explore, Cohort Analyzer, Studies, Data Model, Resources, and About. Below the navigation is a search bar labeled "search C3DC" and a "Search" button. A banner at the top indicates it's an official website of the United States government. The main content area has a teal header with the study title and a "Participants in this Study: 1,119" link. Below this, there's an "Overview" section with a "STUDY DESCRIPTION" paragraph. The paragraph discusses the TARGET NBL dataset, mentioning ~214 fully characterized patient cases with neuroblastoma (all tumor/normal pairs, 10 with relapse sample as well) and some cell lines and xenografts. It also notes 244S cases as well. Each case includes 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 miRNA-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. It encourages users to visit the TARGET website (<http://www.cancer.gov/ccl/research/genome-sequencing/target>) for additional information on this and other TARGET genomics projects. A link to the "TARGET Publication Guidelines" is provided for updated details on the sharing of any TARGET substudy data.

Source Files: TARGET_NBL_ClinicalData_Discovery_20220125.xlsx (1.4 MB)
TARGET_NBL_ClinicalData_Validation_20220125.xlsx (1.4 MB)

Participants: 1,119
Survival Records: 1,119
Diagnoses: 3
Anatomical Sites: 48

Figure 21A: The Studies Details page – downloadable Open Access Source file data for TARGET datasets

The screenshot shows the C3DC Studies Details page for DBGAP ACCESSION: phs001437. The layout is similar to Figure 21A, with a navigation bar, search bar, and study-specific banner. The main content area has a teal header with the study title and a "Participants in this Study: 267" link. Below this, there's an "Overview" section with a "STUDY DESCRIPTION" paragraph. The paragraph describes the Pediatric Preclinical Testing Consortium (PPTC), which 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_v3.1.0.xlsx (1.4 MB)

Participants: 267
Survival Records: 86
Diagnoses: 28
Anatomical Sites: 35

Figure 21B: 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.

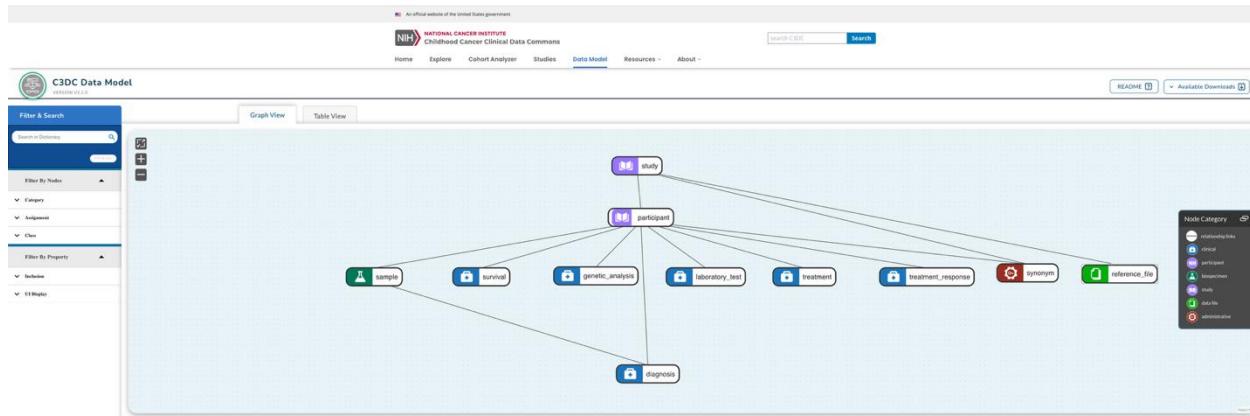


Figure 22A: The Data Model Navigator (DMN) page – Graph with nodes visible and dynamically arranged

This screenshot shows the Table View of the C3DC Data Model Navigator. The top navigation bar and sidebar are identical to Figure 22A. The main content area is a table listing various data types with their descriptions and assignment details. The table rows include:

Category	Description	Assignment	Class
Clinical	The diagnosis node is comprised of properties which describe the participant's disease(s).	Core	Primary
Survival	The survival node is comprised of properties which describe metadata related to vital status and medical events at timepoints for a participant.	Core	Secondary
Genetic Analysis	The genetic analysis node is comprised of properties which describe genetic details of the participants.	Extended	Secondary
Laboratory Test	The laboratory test node is comprised of properties which describe laboratory tests and results for a participant.	Extended	Secondary
Treatment	The treatment node is comprised of properties which describe the treatment regimens and time periods for a participant.	Core	Primary
Treatment Response	The treatment response node is comprised of properties which describe the patient's response to a treatment.	Core	Secondary
Participant	The participant node is comprised of properties which describe the participant attributes.	Core	Primary

Figure 22B: The Data Model Navigator (DMN) page – Table view with readme

C3DC Resource Page

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

An official website of the United States government

NIH NATIONAL CANCER INSTITUTE
Childhood Cancer Clinical Data Commons

Search C3DC

Home Explore Cohort Analyzer Studies Data Model Resources

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/CBII/c3dc-model>

GitHub Harmonization Repository

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

https://github.com/chicagopcde/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 **76,600 CDEs** covering many aspects of cancer research. Access it here: <https://cadrsr.cancer.gov/oncadata/Home.jsp>. There are over 490 pediatric CDEs used by various entities like CCDI Data Ecosystem and Pediatric Clinical Data commons. The **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 23: 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

An official website of the United States government

NATIONAL CANCER INSTITUTE
Childhood Cancer Clinical Data Commons

Home Explore Cohort Analyzer Studies Data Model Resources About

search C3DC Search

About the Childhood Cancer Clinical Data Commons

Childhood Cancer Data Initiative

The NCI's Childhood Cancer Data Initiative (CCDI) [↗](#), which the Childhood Cancer Clinical Data Commons is part of, is an initiative seeking to build a community centered around childhood cancer care and research. Through enhanced data sharing, the initiative works to improve understanding of cancer biology, preventive measures, treatment, quality of life, and survivorship, as well as ensure that the community can learn from every child with cancer. C3DC is part of the CCDI Data Ecosystem. Learn more about other CCDI data and resources on the CCDI Hub. [Sign-up for email updates ↗](#) from NCI about CCDI.

Childhood Cancer Clinical Data Commons

The Childhood Cancer Clinical Data Commons (C3DC) is an open-access web application that serves as the primary source for deidentified, individual-level harmonized data that describes the demographic and phenotypic characteristics of participants. This harmonization process uses a standard data dictionary [↗](#) consisting of Common Data Elements (CDEs). C3DC employs a common data model to facilitate cohort analyses and correlative analytics with data in other datatype-specific commons. The data model has been deposited in [GitHub ↗](#).

The C3DC offers several key features:

- Allows researchers to search for harmonized participant-level clinical data collected from multiple studies.
- Facilitates longitudinal data analyses.
- Enables custom/synthetic cohort creation and data downloading for subsequent local analyses.

Citing the C3DC

NCI expects users to acknowledge CCDI data use as follows:

"The results published here are, in whole or in part, derived from the analysis of data listed in the C3DC ([clinicalcommons.ccdi.cancer.gov](#)), established by the National Cancer Institute's Childhood Cancer Data Initiative (CCDI)."

To cite individual studies, note the CCDI study ID (e.g., phs002790) and include the name and URL or link for the C3DC ([clinicalcommons.ccdi.cancer.gov](#)), along with the phrase, "established by the National Cancer Institute's Childhood Cancer Data Initiative (CCDI)."

Example: "The results analyzed and <published or shown> here are based in whole or in part from analyzing the Molecular Characterization Initiative data listed in the C3DC ([clinicalcommons.ccdi.cancer.gov](#)) under study ID phs002790. The data were accessed from the NCI's Cancer Research Data Commons ([datacommons.cancer.gov](#)). The C3DC was established by the National Cancer Institute's Childhood Cancer Data Initiative (CCDI)."

Questions for C3DC?

The Childhood Cancer Data Initiative (CCDI) welcomes community input to improve this web application usability. Please send your feedback and suggestions to [ncichildhoodcancerdatainitiative@mail.nih.gov ↗](#). Your contributions are valuable to enhancing the user experience.




Figure 24: 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 with a dark teal header featuring a network graph background and the text "C3DC Announcements". Below the header are three cards, each representing a different type of update:

- Datasets Update**: Includes a thumbnail of a circuit board, a date (October 22, 2025), and a brief description: "This release includes newly harmonized data for the following studies: phs003160, phs003161, phs003111, phs001327, phs002620, phs002276, and phs002883, along with updates to several existing CCDI datasets. These new datasets add approximately 3,750 participants." A "Read More" button is present.
- Data Model Update**: Includes a thumbnail of a network graph, a date (October 22, 2025), and a brief description: "The C3DC data model has been enhanced, including the new Laboratory Test node, which captures a multitude of different analyses, and the new Genetic Analysis node that captures gene mutation information for participants." A "Read More" button is present.
- Resource Update**: Includes a thumbnail of a megaphone, a date (October 22, 2025), and a brief description: "We are excited to announce the release of C3DC Application Version 1.7.0! This release incorporates the newly harmonized datasets, along with updates to existing datasets. Key enhancements to the Explore Page and Cohort Analyzer improve the user experience and provide more comprehensive data access. For full release details, please see the Release Note." A "Read More" button is present.

Figure 25: 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.

Data Update

RELEASE VERSIONS

▲ 2025

Data Release 7.0	OCT 22, 2025
Data Release 6.0	JUN 18, 2025
Data Release 5.0	MAR 5, 2025

► 2024

Data Release 7.0

Study Overview

October 22, 2025

Childhood Cancer Survivor Study (CCSS) (New Dataset) - phs001327

- 517 Participants
- 0 Diagnoses

Gabriella Miller Kids First Pediatric Research Program in Pediatric T-Cell Acute Lymphoblastic Leukemia (New Dataset) - phs002276

- 1,358 Participants
- 1 Diagnosis

Feasibility and Clinical Utility of Whole Genome Profiling in Pediatric and Young Adult Cancers (New Dataset) - phs002620

- 113 Participants
- 42 Diagnoses

NCI-COG Pediatric MATCH Precision Medicine Clinical Trial (New Dataset) - phs002883

- 1,239 Participants
- 136 Diagnoses

Clonal Evolution During Metastatic Spread in High-Risk Neuroblastoma (New Dataset) - phs003111

- 129 Participants
- 0 Diagnoses

Figure 26: 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.

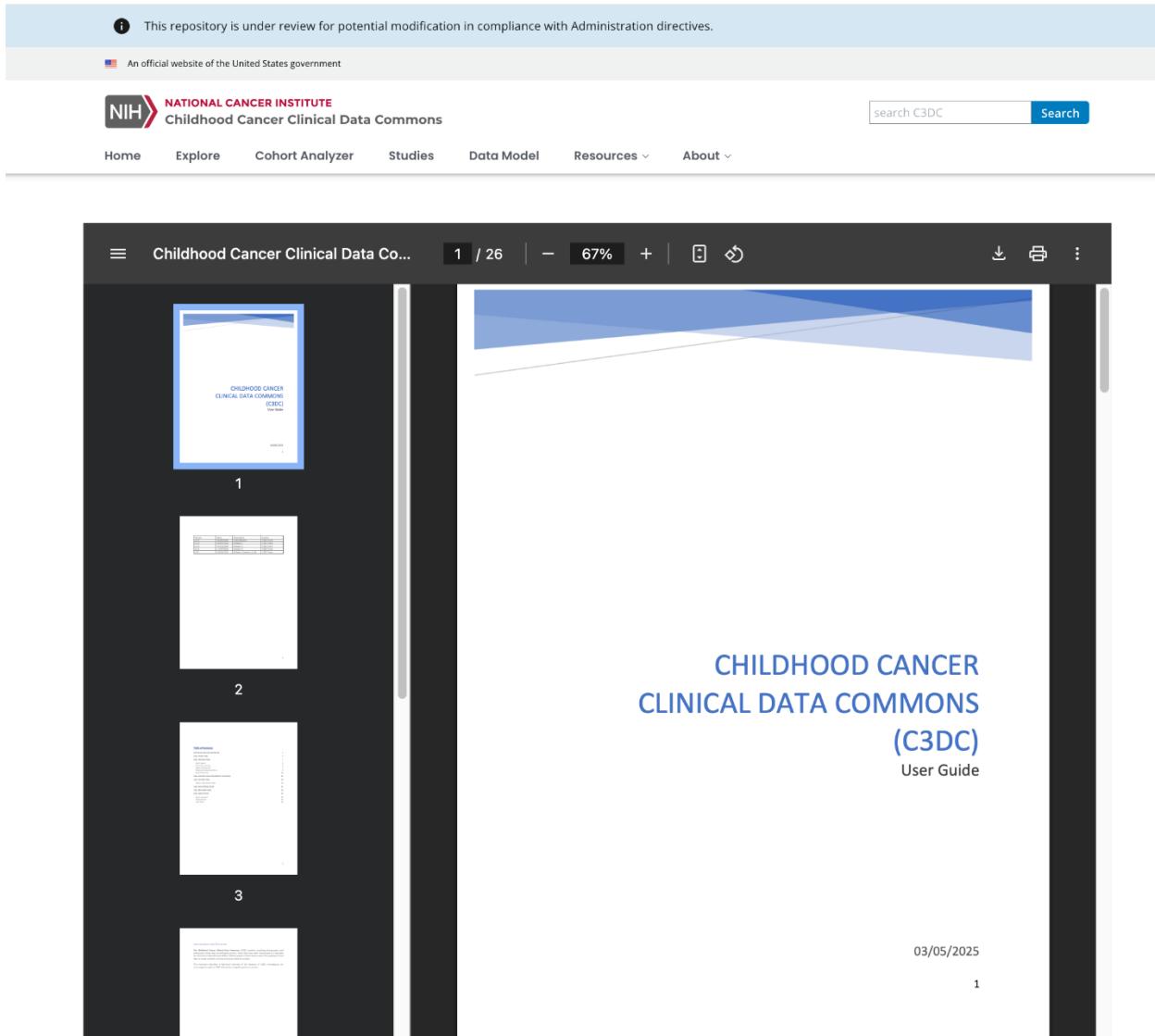


Figure 27: The User Guide