CHILDHOOD CANCER DATA CATALOG (CCDC)

User Guide

2/12/2025

Version	Date	Description	Author
1.0.0	04/15/2022	Initial Version	CCDC Team
1.1.0	05/15/2022	Updated information about result filtering and display. Added resources to the site. Replaced some images with improved contrast versions.	CCDC Team
1.2.0	9/15/2022	Updated Home page section for Featured Items. Updated Search Catalog for search changes. Added Site Update page section. Updated screenshots for site changes. Updated Appendix A with datasets updates.	CCDC Team
1.2.1	10/18/22	Updated Appendix A	CCDC Team
1.2.2	11/17/22	Updated Site Update Page section to include information on viewing entries. Updated Appendix A	CCDC Team
1.3.0	1/11/2023	Updated the Dataset Details page to include additional available Core Data Elements and added an additional section defining Charts. Updated Resource Types to include Analytic Tools. Updated Appendix A	CCDC Team
1.3.1	2/15/2023	Updated lists of Core Data Elements and Data Content Types. Updated Appendix A	CCDC Team
1.3.2	3/15/2023	Added new section for CCDI Datasets menu link. Update screenshots for site change. Updated Appendix A	CCDC Team
1.3.3	4/19/2023	Updated Dataset Additional Data Elements. Updated Resource Type information. Updated Appendix A	CCDC Team
1.3.4	5/17/2023	Updated Dataset Additional Data Elements. Updated Appendix A	CCDC Team
1.3.5	6/14/2023	Updated Appendix A	CCDC Team
1.3.6	7/19/2023	Updated the section for accessing CCDI data from the menu bar and added a section on the CCDI data access guide. Updated alternate text of some document images. Updated Appendix A	CCDC Team
1.3.7	8/16/2023	Added Biorepository Resource Type and Biospecimens Data Content Type. Updated Repository Resource Type to be Data Repository. Updated Appendix A	CCDC Team
1.3.8	9/20/2023	Updated Appendix A	CCDC Team
1.4.0	10/19/2023	Updated Appendix A Updated Table of Contents	CCDC Team
1.4.1	11/15/2023	Updated Appendix A	CCDC Team

1.4.2	12/13/2023	Updated Appendix A	CCDC Team
1.4.3	1/17/2024	Updated Appendix A	CCDC Team
1.4.4	2/14/2024	Updated Appendix A	CCDC Team
1.4.5	3/13/2024	Updated Appendix A	CCDC Team
1.4.6	4/10/2024	Added new section for CCDI Supported Applications. Expanded Contribute to the CCDI Data Catalog section. Updated figure numbering. Updated Appendix A	CCDC Team
1.4.7	5/17/2024	Updated Appendix A	CCDC Team
1.4.8	6/12/2024	Updated Appendix A – CLIC removed	CCDC Team
1.4.9	8/21/2024	Updated Appendix A	CCDC Team
1.5.0	12/11/2024	Added 2 new resources: Inventory of Division of Cancer and Epidemiology and Genetics (DCEG) Research and Specimen Resource Locator (SRL). Added new features: support direct download of Data Digest files and added TARGET data file links. Updated Appendix A.	CCDC Team
1.5.1	2/12/2025	Added resource NCTN Biobanks. Updated figures to show new headers, updated counts and generally bring user guide more current. Updated Appendix A.	CCDC Team

Contents

Introduction and Overview 5
Home page 5
Search Catalog Page 6
Dataset Results Views 7
Card View 8
Table View 8
Search 9
Export 10
Dataset Detail Page 10
Administrative data 11
Metadata 11
Core Data Elements 12
Additional Data Elements 13
Charts 13
Participating Resource Page 14
Resource Filtering 15
Resource Card View 16
Card View 16
Participating Resource Detail Page 18
Administrative data 19
Descriptive Data 19
Resource Description 20
Resource Tools 20
Data Content Type 20
Data Access 20
Dataset Summaries 20
CCDI Resource and Datasets 22
CCDI-Supported Applications 22
Accessing CCDI Data 23
About CCDI Data Catalog Page 23
Contributing to the CCDI Data Catalog 24
Glossary Page 25

Site Update Page 26
User Guide 27
Documentation Search Page 28
Additional Support 29
Appendix A: Participating Resources 30

Introduction and Overview

The Childhood Cancer Data Catalog (CCDC) is an inventory of pediatric oncology data resources developed as part of the National Cancer Institute's (NCI) Childhood Cancer Data Initiative (CCDI). The CCDC includes listing of childhood cancer repositories, registries, programs, knowledgebases, as well as other catalogs that either manage or refer to pediatric cancer data. The aim of the CCDC is to help researchers, clinicians, and citizen scientists learn about existing pediatric resources for the development of new biomedical hypotheses, facilitate new secondary analyses of data within existing resources, and to provide connections to existing pediatric oncology research sites.

This document describes the purpose and the general use of the CCDC application. By design, the CCDC only collects and holds the metadata from a specific resource and not the primary participant or sample data of that resource. Information to help understand this metadata is defined below, but users interested in specific resource data are encouraged to reach out to content holders using the point of contact (POC) information available on the CCDC site.

Home page

The CCDC Home page is the landing page for users navigating to the CCDC site. It provides a brief introduction to the site, allows users to perform a search of the cataloged data resources, shares information about updates to the site, and shares a rotating carousel of participating resources available on the site (Figure 1). This page is also reachable by selecting 'Home' on the menu bar from any page in the site.

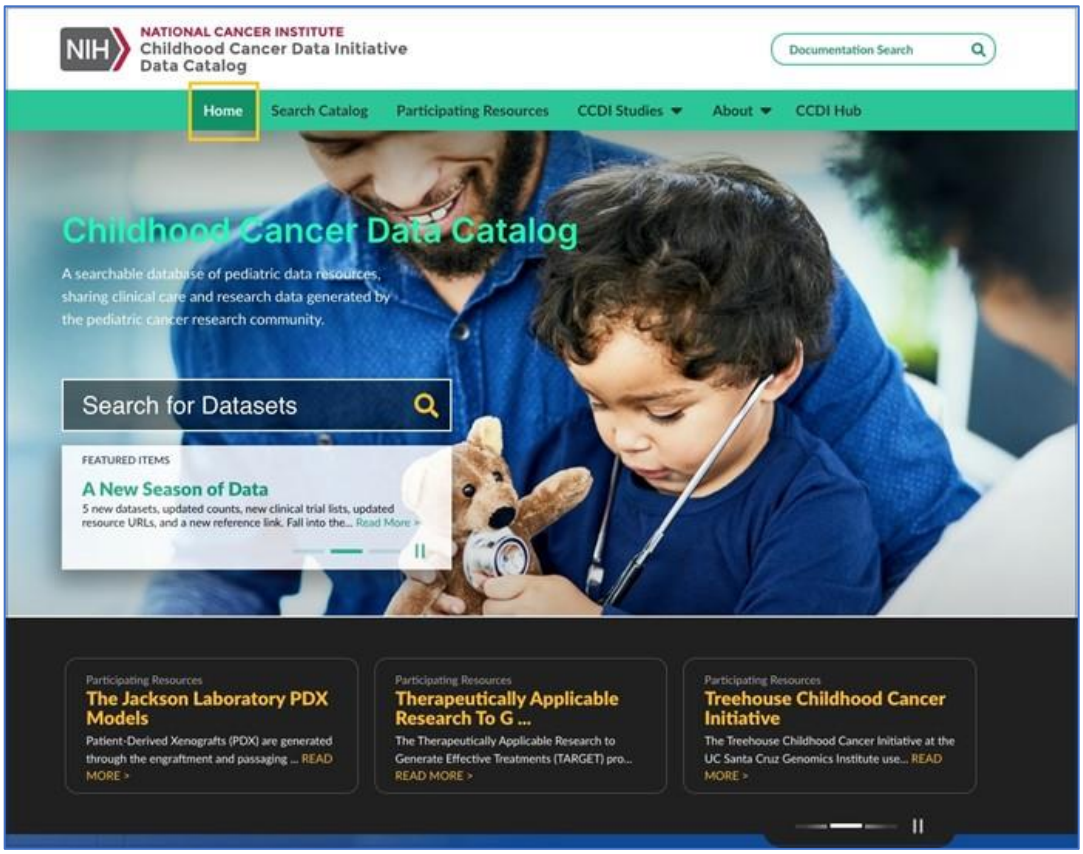

Figure 1: CCDC Home Page

Users wishing to search the cataloged data resources can use the “Search for Datasets” field (Fig. 1) to perform a search that will automatically navigate a user to the Search Catalog page. More information on the Search Catalog page and the catalog search parameters is available below.

Below the search field is a Featured Items widget. This widget highlights site changes and shares information related to the CCDI. The featured items widget rotates between three featured items that may highlight updates to the CCDC site or are informational in nature. Featured items may include links that navigate within the site or to external sites related to CCDI.

The home page also features a rotating carousel of participating resources available on the site to which users can navigate. This carousel contains example resources that will periodically change. Each example lists the resource name and a brief introduction to the resource. Users can either select the resource name or the ‘READ MORE >’ links to navigate to the resource’s details page to learn more about the resource. More information about the participating resource details page is available below.

Search Catalog Page

The CCDC Search Catalog page allows users to search the individual datasets available from CCDC participating resources (Figure 2). You can navigate to the page by selecting ‘Search Catalog’ on the menu bar. Introductory information is available for each dataset along with a link to a more detailed dataset page. The Search Catalog page features a free text search that allows for filtering datasets.

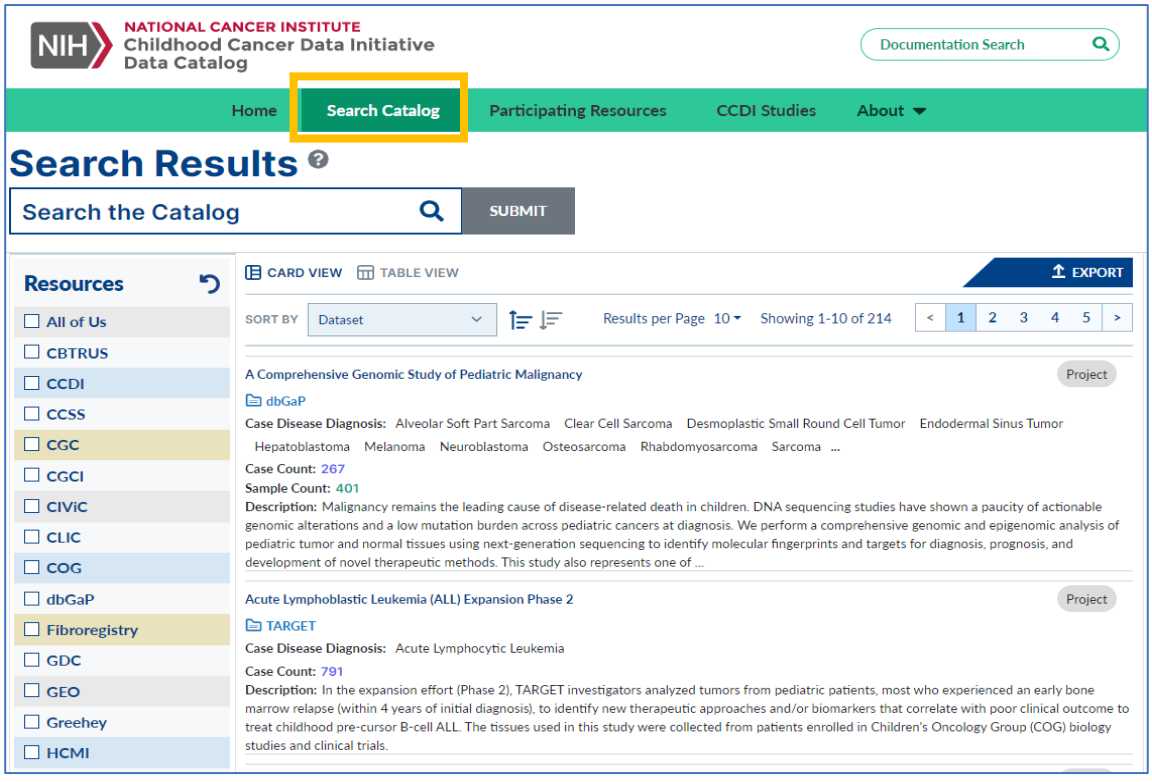

Figure 2: CCDC Search Catalog Page

Please note, the datasets available on the site are periodically updated and the types of datasets available may change.

Dataset Results Views

Information about available datasets is displayed in the results pane on the right-hand side of the Search Catalog page. The results pane can be configured in two ways with either a Card View (Figure 3) or a Table View (Figure 4). Each view can be toggled back and forth without altering the results of a performed search.

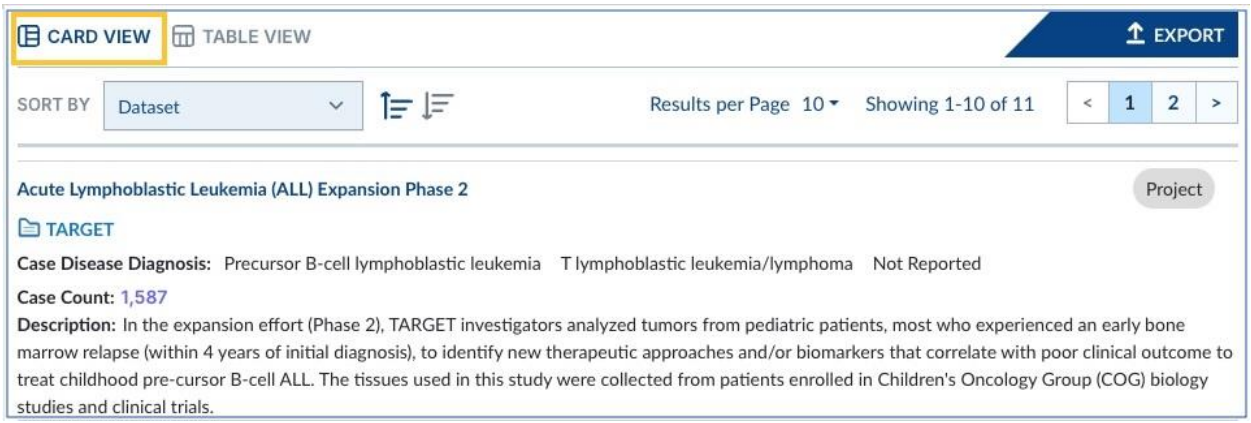

Figure 3: Search Catalog Results Pane Card View

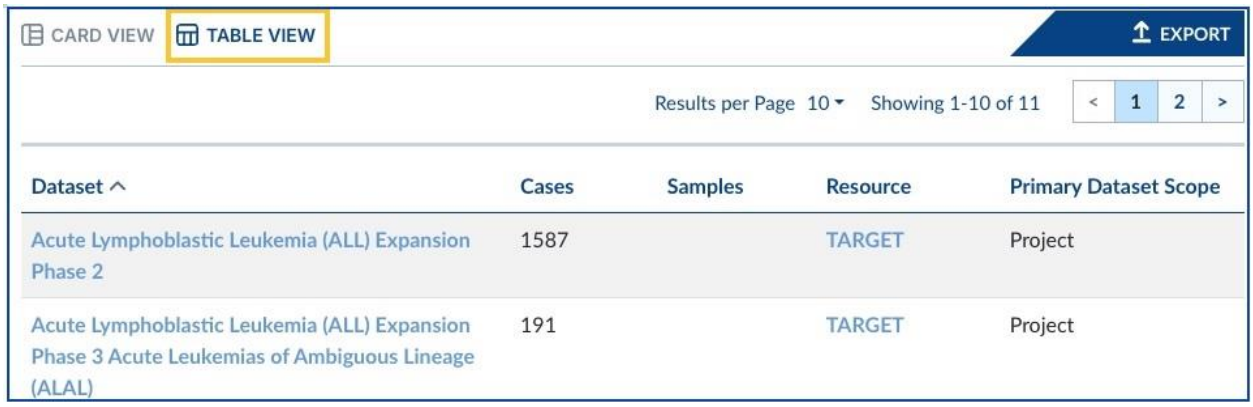

Figure 4: Search Catalog Results Pane Table View

In both views, users can select a dataset name to navigate to a page with more detailed information about the dataset. Within each result, selecting the participating resource name/abbreviation (i.e., Kids First) will navigate to a more detailed page about that participating resource.

The information in the results pane from the Card View can be sorted with the 'Sort By' filter (Figure 5). Filtered results are displayed in ascending order by default but can be changed by selecting the descending order icon. In the Table View, table columns can be sorted by clicking the column headers.

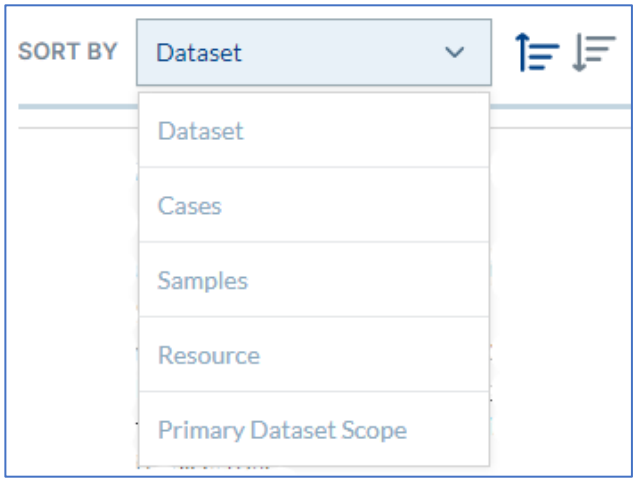

Figure 5: Search Catalog 'Sort By' Filter

The result pane displays 10 search results by default. If a search returns more than 10 results, the subsequent results are available for review on subsequent pages. These pages can be accessed using the '<' or '>' buttons or a specific page number on the pagination bar at the top and bottom of the results pane (Figure 6). If returned results contain fewer than 10 results, the pagination buttons are disabled. Users can select the 'Results per Page' dropdown to modify the number of displayed results.

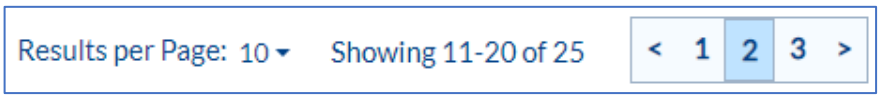

Figure 6: Search Catalog Pagination Bar

Card View

The Card View (Figure 3) is meant to provide a brief highlight of the datasets available in the catalog. Each card provides some abbreviation of selected metadata about a dataset for an initial assessment of relevancy. In the Card View, information for the following data elements is available for display:

- **Dataset Name**
- **Participating Resource Name**
- **Case Disease Diagnosis**
- **Case Count**
- **Sample Assay Method**
- **Sample Count**
- **Description**
- **Primary Dataset Scope**

If a dataset does not have information for a data element, the data element will not be displayed. The card view may also display '**Other Match**' and an additional data element if information from the search text is contained within another dataset data element. 'Other Match' information will only appear when performing a search.

Table View

The Table View (Figure 4) is meant to provide a selection of information on the datasets available in the catalog. The Table View displays the following name and count data elements:

- **Dataset Name**
- **Case Count**
- **Sample Count**
- **Participating Resource Name**
- **Primary Dataset Scope**

Search the Catalog

There is a text field at the top of the page (Figure 7) to search the catalog datasets. The search will perform a full and partial word search for any terms or text entered that meet the search parameters. The parameters, defined below, are also available on the site by selecting the '?' tooltip next to the search field.

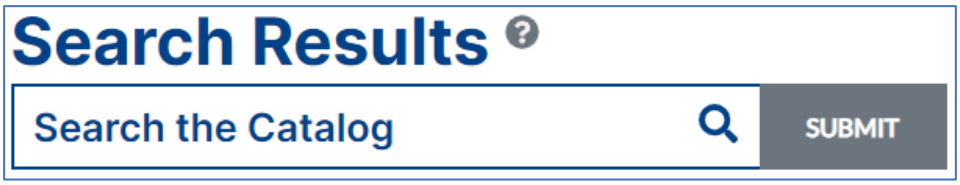

Figure 7: Search Catalog Search Field

Parameters for searching the catalog:

- A minimum of 3 characters are needed for a term search.
- Searches will return both full and partial word results (i.e., 'leuk' will return results for leukemia).
- If you search multiple terms (i.e., lymphocytic survivors) the search results will return sources that contain all the specified terms (AND Boolean operator).
- Results can be filtered by Participating Resources by checking a resource in the resource column. Selecting multiple Resources will filter as an OR Boolean operator.
- Anatomical site searches leverage the NCI Thesaurus to display Case Tumor Site synonym matches. For example, a search for 'eye' will also return results for 'orbit.'

To execute a search, enter the text of interest and either select the 'Submit' button to the right of the search field or hit enter/return on your keyboard. This will execute a search and you should see a prompt with the searched text appearing to the right of the search field as a 'Search Text' bubble (Figure 8). With an executed search, the results pane will be updated to display the datasets that contain a match for the defined text. To clear the search field, select the 'X' on the right side (Figure 8). Clearing the field will remove typed text but does not reset an executed search.

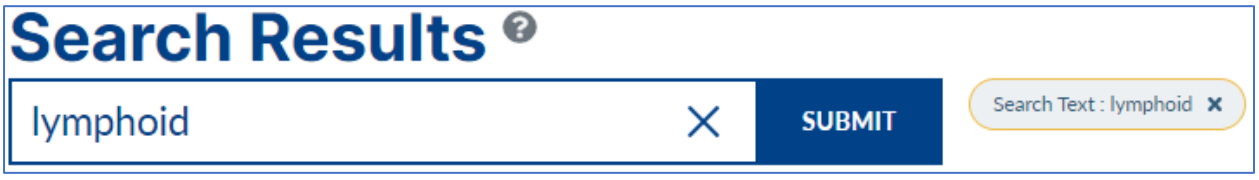

Figure 8: Search Catalog Valid Search Execution

Additionally, the Resource column, to the left of the results pane, will update to display the participating resources that have datasets that match the search. To reset the search, you can select the 'x' within the Search Text bubble.

The catalog search can also be filtered by selecting a resource (i.e., CCSS) from the Resource column (Figure 9). Selecting a Resource will automatically populate a prompt with the searched text appearing to the right of the search field as a 'Resources' bubble (Figure 9). A Resource filter can be applied alone or in conjunction with a text search. To reset the search, you can select the 'x' within the Resource bubble or select the circular back arrow in the Resource column on the left.

Figure 9: Search Catalog Resource Pane Filter

Export

The Search Results page features an export that allows the results of a search to be exported for local manipulation or record keeping. To export results, there is an 'Export' link in the upper right of the results. This link will create an export of the results shown in the results pane. The file is exported as a comma-separated values (.csv) file that can be opened in a local text or spreadsheet tool.

The file contains the information and data elements for a dataset that has been gathered for the CCDC. Each row is an individual dataset and includes information not found on the catalog search page. Please note that the information found on each row is unique to each dataset and that datasets, even within a single participating resource, may contain varying amounts of content per data element.

Dataset Detail Page

The Dataset Details page (Figure 10) provides detailed information about that dataset. The Dataset Details page is available from either the Search Catalog page or from a participating resources details page. The Dataset Details page contains a listing of all metadata the CCDC site has collected about a dataset. The page is broken up into two parts, a header with administrative information and a body that contains count and summary data.

The data on the Dataset Details page is only updated periodically and data on the CCDC site may be out of sync with participating resource's own data sharing site. For the most up to date information about a dataset, it is recommended that you verify any data of interest with the resource owner's website.

Figure 10: Dataset Details Page

Administrative data

At the top of each Dataset Details page is an administrative header (Figure 11). The administrative header contains the name of the dataset, a link to the participating 'Data Resource' on the CCDC site to which the dataset belongs, and Point of Contact information for the dataset. When possible, the point of contact contains both a name and an email address for a resource owner.

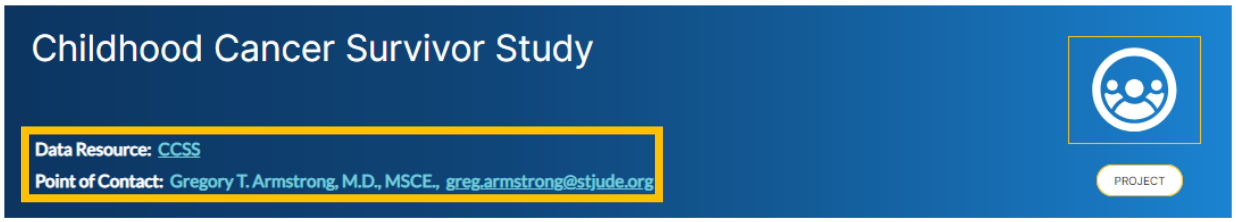

Figure 11: Dataset Details Page Administrative Data

Metadata

The main body of the Dataset Details page contains the collected dataset metadata (Figure 12). In addition to the dataset description, the dataset metadata describes the 'Core Data Element' and the 'Additional Data Elements' of the dataset. These data elements are designed to provide a high-level

overview of the data available within the dataset and provide supporting information about the provenance of the dataset. A visualiation of the Core Data Elements is available under Charts.

For both Core and Additional Data Elements, if information was not available for a data element, that data element will not appear on the Dataset Details page. You should expect some variation in available data between datasets.

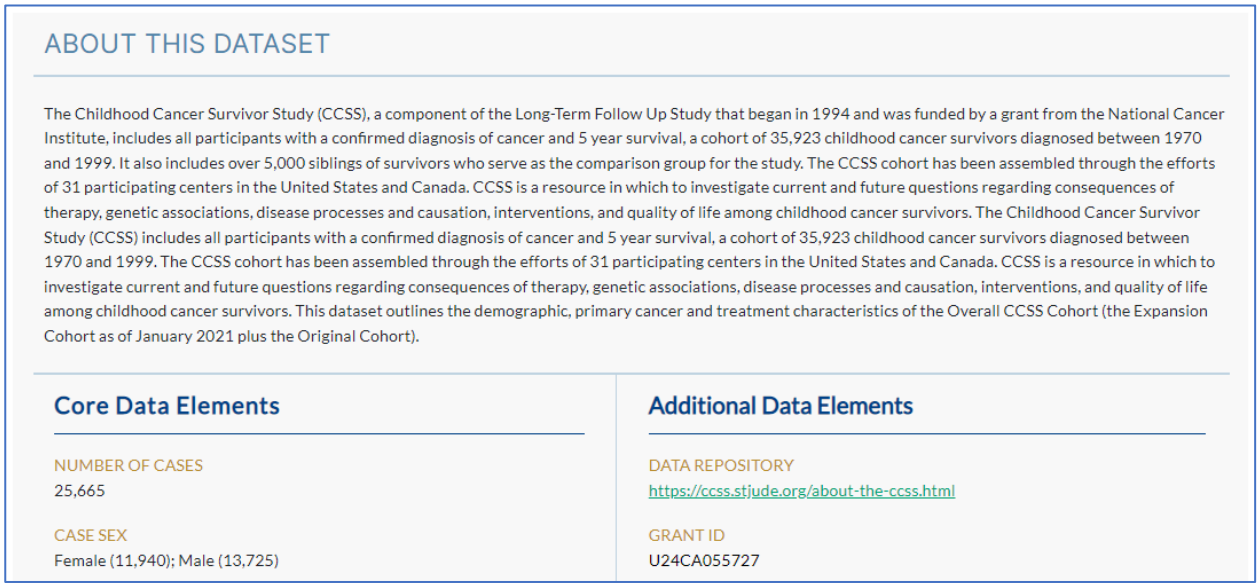

Figure 12: Dataset Details Page Metadata

Core Data Elements

Core data elements describe and provide count information about data within the dataset. These data elements describe the participants, related disease information, and information about sample, if samples are part of the dataset. For each numeric data elements, such as Number of Cases, the numeric count value is of that element displayed. For response type data elements, such as Case Sex, the response is displayed along with the count for that response (i.e., Female (11,940)).

The following data elements are displayed as the core data elements:

- **Number of Cases**
- **Case Sex**
- **Case Age at Diagnosis**
- **Case Race**
- **Case Ethnicity**
- **Case Disease Diagnosis**
- **Case Tumor Site**
- **Case Treatment Administered**
- **Case Treatment Outcome**
- **Case Proband**
- **Number of Samples**
- **Sample Assay Method**
- **Sample Analyte Type**

- **Sample Anatomic Site**
- **Sample Composition**
- **Sample in Normal**
- **Sample is Xenograft**

Additional Data Elements

Additional Data Elements are generally the data elements that provide supporting information about a dataset such as grant identifier. However, some datasets do include additional information that describes the data within a dataset, such as the data element 'Has Mutation Data'. Select datasets may also have links to files in other repositories as indicated in the 'Available Files' element.

The following are the standard data elements may be displayed in the additional data element column:

- **dbGaP Study Identifier**
- **Available Files**
- **Clinical Trial Identifier**
- **Clinical Trial List**
- **Data Repository**
- **GEO Study Identifier**
- **Grant Information**
- **Projects**
- **Program ID**
- **Published In**
- **SRA Identifier**

Charts

The Charts section provides an alternate, visualized presentation of the core data elements. Charts will display each core data element available for a dataset, except Number of Cases and Number of Samples which will not be displayed. The data elements displayed will be shown as a donut chart except Case Age at Diagnosis. Case Age at Diagnosis will display as a Histogram. By default, the first available core data element will be displayed. The displayed information can be changed by selecting the dropdown depicted as 'B" in the charts below (Figure 13).

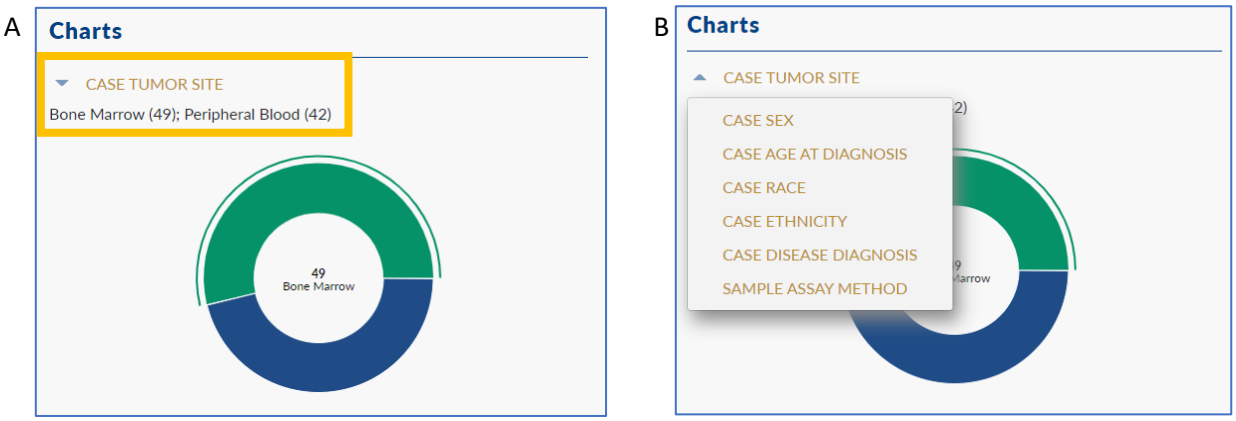

Figure 13: Dataset Details Page Charts. A – Dropdown Menu Closed B – Dropdown Menu Open

Participating Resource Page

The CCDC Participating Resource page allows users to view and filter the resources that CCDC datasets are drawn from (Figure 14). The Participating Resources page can be navigated to by selecting 'Participating Resource' on the menu bar. The resource page includes introductory information for each resource available on the site along with a link to view a more detailed resource page. Resources on the page can be filtered by 'Resource Type' and by 'Data Content Type' to view specific resources.

Figure 14: Participation Resource Page

The catalog currently includes the following types of resources.

- **Analytic Tool** – Any platform, methodology, framework or other software designed for the use of and interpretation of biomedical research data.
- **Biorepository** – A biorepository is a facility that acts as a library for biospecimens, allowing the biospecimens to be available for use in future research. A biospecimen may be from people, animals, or other living organisms. A biorepository will be involved in collecting, cataloguing, and storing biospecimens. The biorepository will also be involved in managing access to and

distributing biospecimens to researchers. Some biorepositories store medical information associated with biospecimens.

- **Catalog** – A data catalog is not a data repository but rather a place where data is described with an index to what is available. A collection of digests and references (e.g., URL or POC) to corresponding research artifacts. There is a consistent structure across the collection of digests to facilitate filtering and identifying research artifacts of interest. A catalog contains some combination of Summary Research Data, Summary Clinical Data, Data Overview, and Resource Metadata.
- **Data Repository** – Biomedical data repositories store, organize, validate, archive, preserve, and distribute data, in compliance with the FAIR Data Principles. It is a system for storing multiple research artifacts, provided at least some of the research artifacts contain Individual Research Data. A data repository often contains artifacts from multiple studies. Some data repositories accept research datasets irrespective of the structure of those datasets; other data repositories require all research datasets to conform to a standard reference mode.
- **Knowledgebase** – Biomedical knowledgebases extract, accumulate, organize, annotate, and link the growing body of information that is related to and relies on core datasets.
- **Program** – A coherent assembly of plans, project activities, and supporting resources contained within an administrative framework, the purpose of which is to implement an organization's mission or some specific program-related aspect of that mission.
- **Project** – Any specifically defined piece of work that is undertaken or attempted to meet the goals of a program and that involves one or more case studies. Also known as a Study or Trial.
- **Registry** – A cancer registry is an information system designed for the collection, storage, and management of data on persons with cancer. An inventory of individuals or samples, usually focused on a specific diagnosis or condition. In some cases, public health laws require collecting information in registries about individuals who have a specific disease or condition. In other cases, individuals provide information about themselves to these registries voluntarily. Thus, a registry contains Individual Clinical Data, but not Individual Research Data

Resources available on the site are periodically updated and the types of resources available may change. A complete listing of current resources is available in Appendix A.

Resource Filtering

To filter the catalog resources there is a collapsible filter menu on left-hand side of the page (Figure 15). Within the filter menu there are two available filters– Resource Type and Data Content Type. For either filter, a selection can be made by checking the boxes to the left of the filter option. There is no limit on the number of filter selections that can be made.

When filtering, if selecting multiple options within a single filter the returned results will show using the OR operator for the selected options (Figure 15B). In this example, the results will display all catalog and program resources in the CCDC. If selecting multiple options across the two filters, the results will return with the AND operator for the selected options (Figure 15C). In this example, the results will display only catalogs with clinical data. To reset the filters either un-select any selected options or refresh the page.

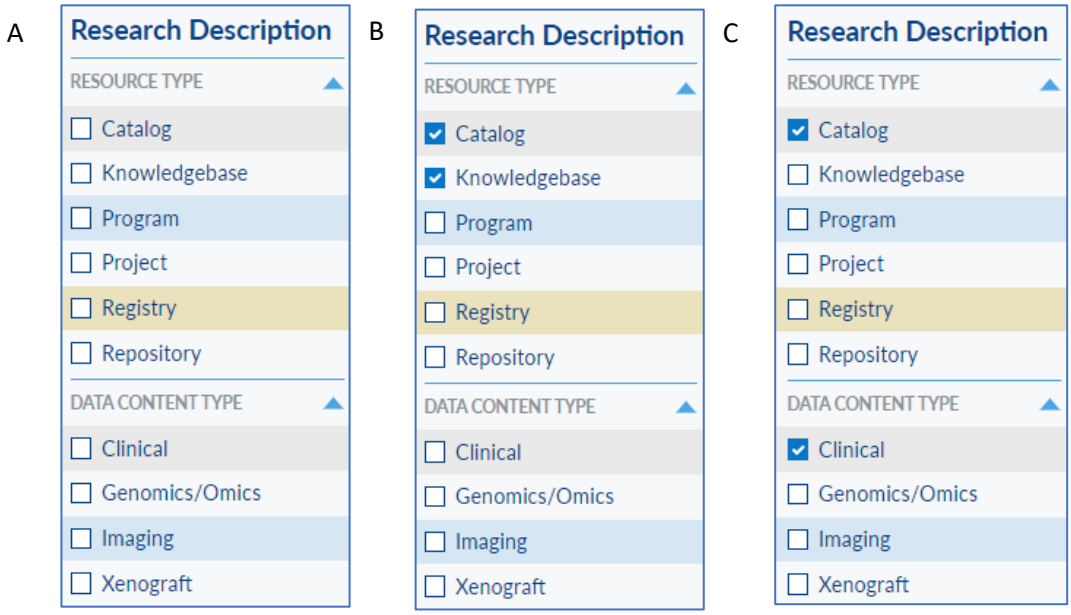

Figure 15: Participating Resource Research Filter. A – No selections. B – Single filter Selection. C – Multi-filter Selection

Resource Card View

Information about participating resources is displayed in the results pane on the right-hand side of the Participating Resource page (Figure 14). The resources in the pane are displayed in a Card View. Within the resource card, select a resource's name to navigate to a page with more detailed information about the resource and its associated datasets.

The result pane displays 10 search results by default. If a filter returns more than 10 results, the results are available for review on subsequent pages. These pages can be accessed using the '<' or '>' buttons or a specific page number on the pagination bar at the top and bottom of the results pane (Figure 14). If returned results contain fewer than 10 results, the pagination buttons are disabled. Users can select the 'Results per Page' dropdown to modify the number of displayed results.

Card View

The resource card view (Figure 16) provides a brief highlight of each resource available in the catalog. Each card provides a summary of the resource, information about connecting directly with a resource, and information about the number of datasets from a resource listed on the CCDC site. Each card displays the following data elements for each resource:

- **Resource Name**
- **About this Resource**

- **Point of Contact Name**
- **Resource URL**
- **Resource Type**
- **Number of available Datasets**

Each resource card links to a page with additional detailed information about a participating resource. Selecting the 'Resource Name' will navigate to the top of the Participating Resource Details page. Selecting the 'Dataset Summaries' will navigate directly to the dataset summaries section of the Participating Resource Details page.

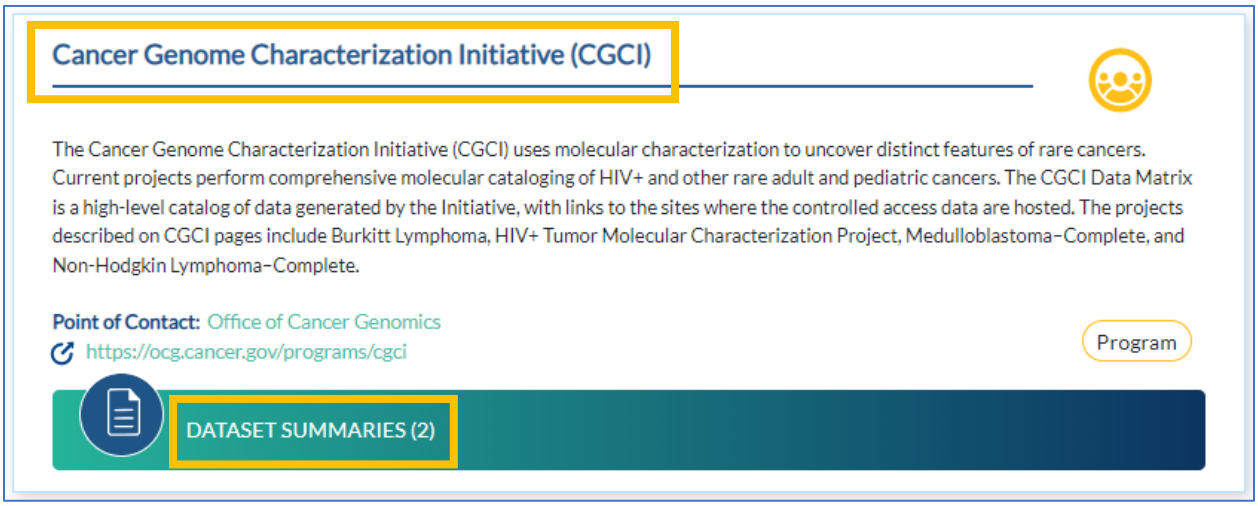

Figure 16: Participating Resource Card View

Participating Resource Detail Page

For each participating resource within the CCDC, there is a page of detailed information about that resource, named the Participating Resource Details page (Figure 17). The resource details can be navigated to by selecting a resource on either the Participating Resource page, the Search Catalog page, or the Dataset Details page. The Participating Resource Details page contains a listing of metadata the CCDC site has collected about a resource. The page is broken up into three parts, a header with administrative information, a body section of descriptive information about the resource, and a body section listing the datasets associated with a resource on the CCDC site.

The data on the Participating Resource Details page is only updated periodically and data on the CCDC site may not be as current as would be found on a participating resource's own data sharing site. For the most up to date information, it is recommended that you verify the data with a resource owner.

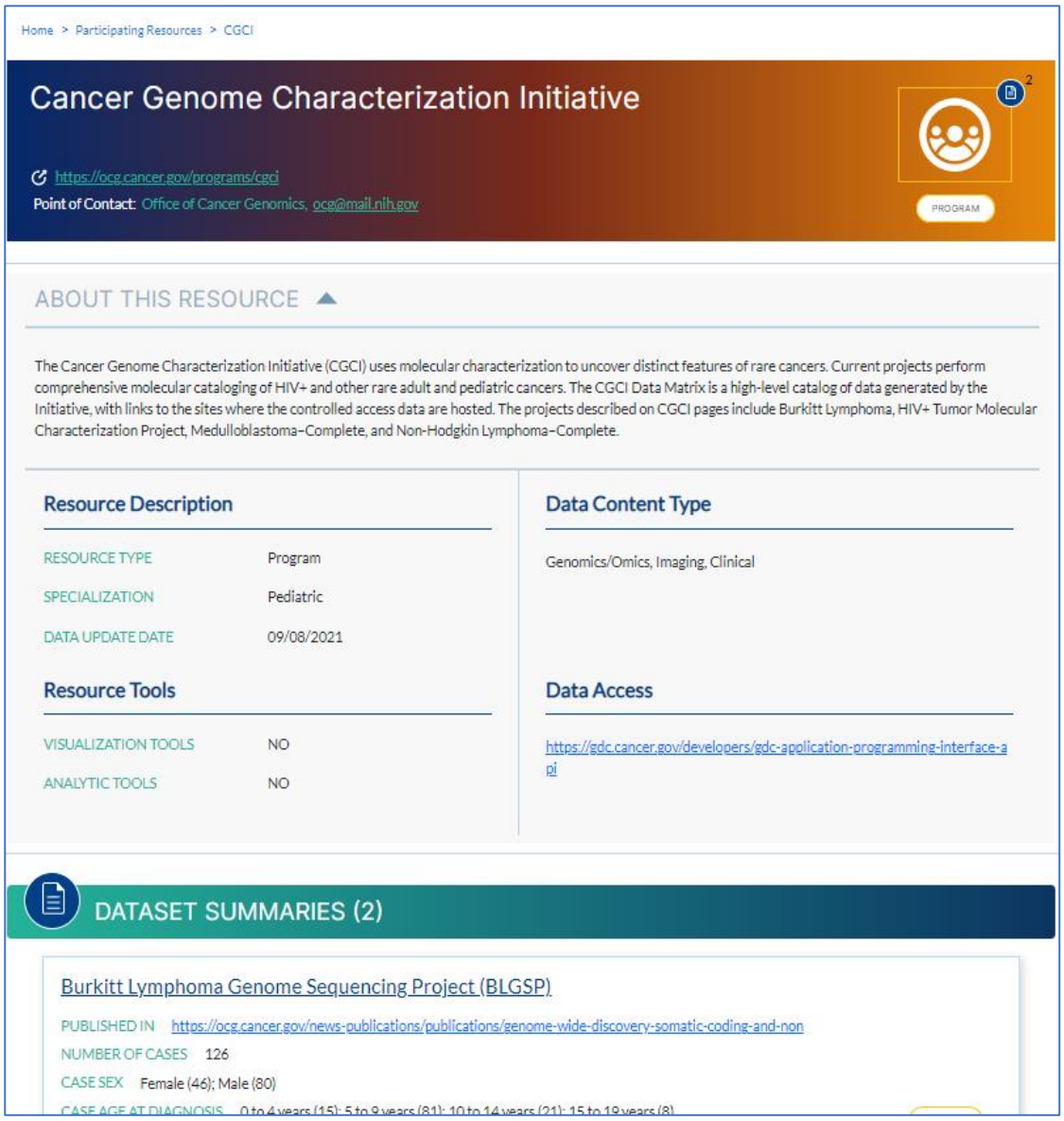

Figure 17: Participating Resource Detail Page

Administrative data

At the top of each Participating Resource Details page is an administrative header (Figure 18). The administrative header contains the name of the resources, a link to the resource's own data sharing website, Point of Contact information for the dataset, and the resource type defined by the CCDC. Point of contact information contains both a name and an email address for a resource owner whenever possible. The header also contains an icon in the upper providing a quick indication of the number of datasets assocated with a resource.

Figure 18: Participation Resource Details Administrative Data

Descriptive Data

The first section of the body of the Participating Resource Details page contains descriptive information about the resource (Figure 19). In addition to the 'About This Resource' description there are four sections of additional information – Resource Description, Resource Tools, Data Content Type, and Data Access.

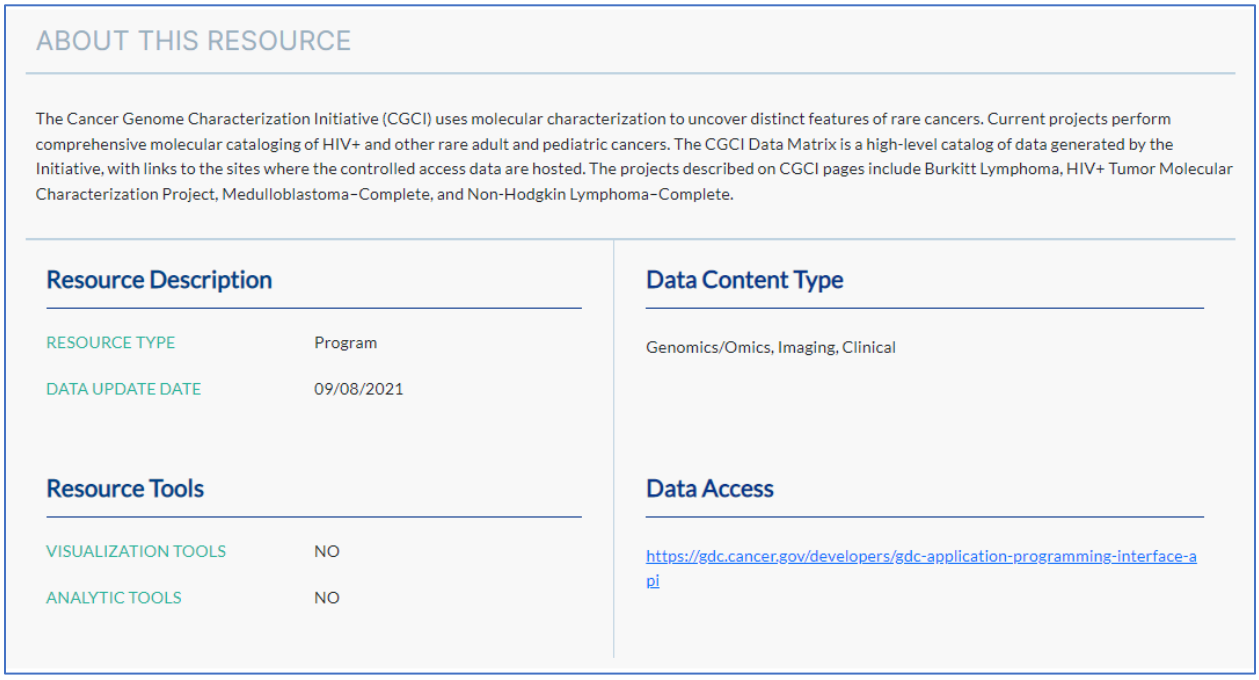

Figure 19: Participating Resource Details Descriptive Data

Resource Description

The Resource Description provides information about how the CCDC organizes the resource and when information on the CCDC site was last updated.

- **Resource Type** – CCDC categorization
- **Data Update Date** – Date the resource was last updated on the CCDC

Resource Tools

The Resource Tools provides information on if a participating resource's own data sharing site includes tools for additional dataset assessment. The CCDC currently identifies these types to assessment tools:

- **Visualization Tools** – A tool for visualization of available data
- **Analytic Tools** – A tool for analytic manipulation of available data

Data Content Type

The Data Content Type defines the type of data within a resource. The CCDC currently identifies these following content types:

- **Biospecimen**
- **Cell Lines**
- **Clinical**
- **Epidemiologic**
- **Genomics/Omics**
- **Imaging**
- **Xenograft**

Data Access

The Data access section provides a link to the resource's own data sharing site where information or datasets are available.

Dataset Summaries

The second section of the body of the Participating Resource Details page contains a listing of the datasets, within the CCDC, associated with a participating resource (Figure 20). All participating resources will have at least one associated dataset and the number of associated datasets is identified within the section header. Each dataset is listed on a card view within the summary.

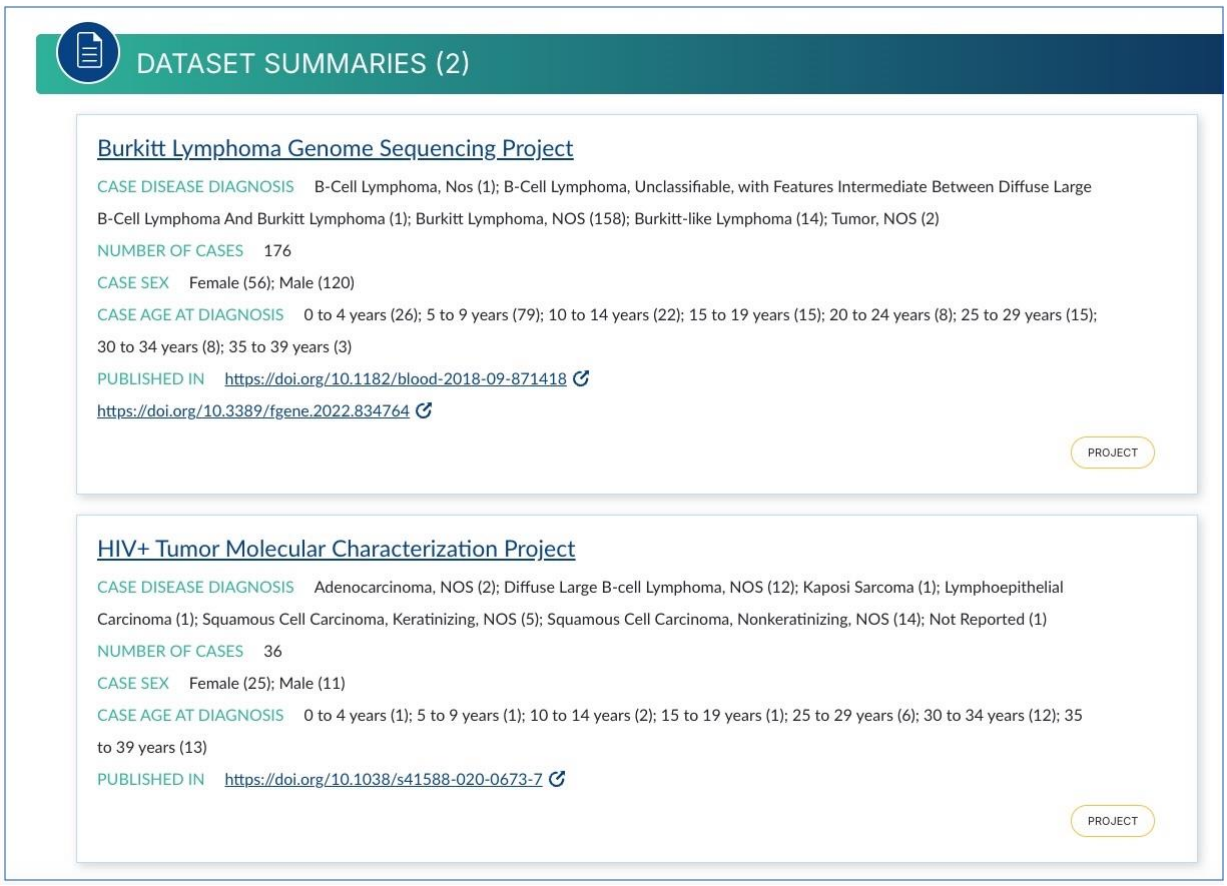

Figure 20: Participating Resource Details, Dataset Summaries

As the Participating Resource Details page already provides background information about the project, the individual dataset cards are more concise and quantitative than those found on the Search Catalog page. The dataset card views display the following data elements for each dataset if available:

- **Dataset Name**
- **Case Disease Diagnosis**
- **Number of Cases**
- **Case Sex**
- **Case Age at Diagnosis**
- **Published In**

Selecting the Dataset Name will navigate to the datasets Dataset Details page where you can learn more about a dataset. Datasets available on the site are periodically updated and the types of datasets available may change.

CCDI Resource and Datasets

Childhood Cancer Data Initiative participating resource details page (Figure 21) provides information about the CCDI and its available datasets. A direct link to the CCDI resource page and the CCDI related datasets is available by navigating to the 'CCDI Resource & Datasets' submenu item within the 'CCDI

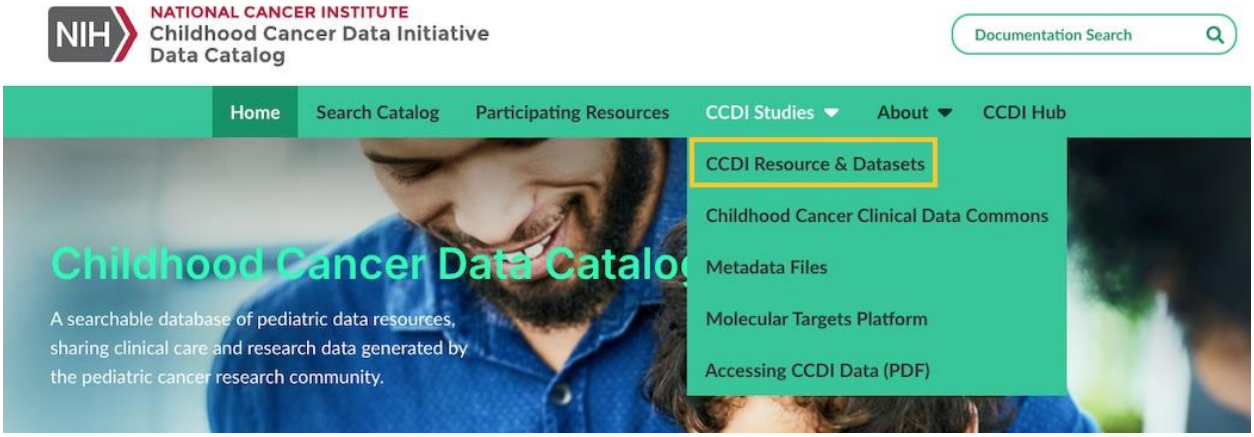

Figure 21: Childhood Cancer Data Initiative Page Selected for CCDI Studies

Studies' dropdown on the menu bar.

Childhood Cancer Clinical Data Commons link-out

The 'CCDI Studies' menu includes a direct link to this CCDI Applications (Figure 22):

- **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.

These links connect researchers, data scientists, and citizen scientists looking to use and connect with CCDI-supported data, tools, and applications.

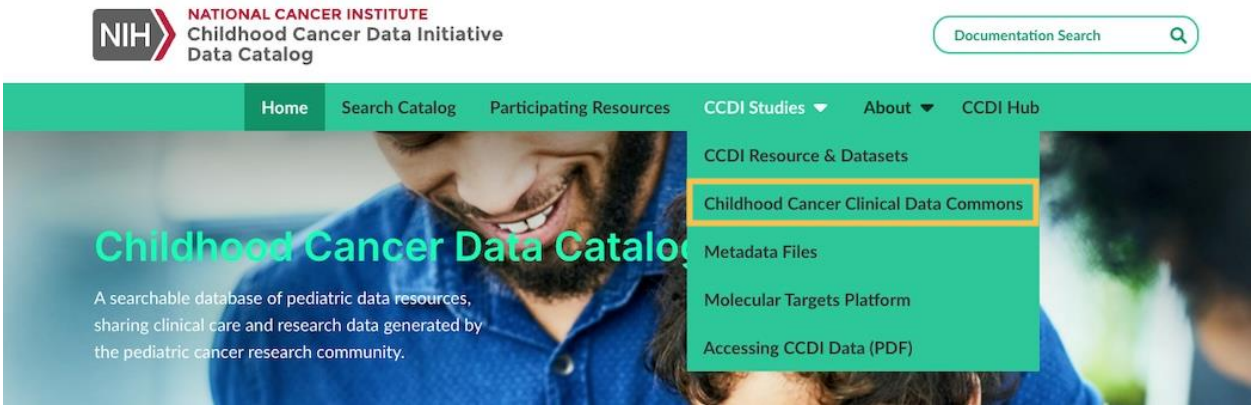

Figure 22: Link to Childhood Cancer Clinical Data Commons

Metadata files Page

Users are now able to download all dataset digests from the user interface menu. Clicking on the "Metadata Files" will produce a "digest_v#.#.#.zip file which will contain all digest files used to update the Resources' information (Figure 23).

Figure 23: Metadata Files Page

Accessing CCDI Data

The 'Accessing CCDI Data' link provides a link to a PDF guide for accessing CCDI data stored at the NCI's Cancer Data Service (CDS) (Figure 24). This guide provides information about how to access, query, and process data from the CCDI. Selecting the guide link will open the guide in a new tab. Please note, for Safari user, clicking on the "User Guide" menu item will directly download the PDF format of the user guide.

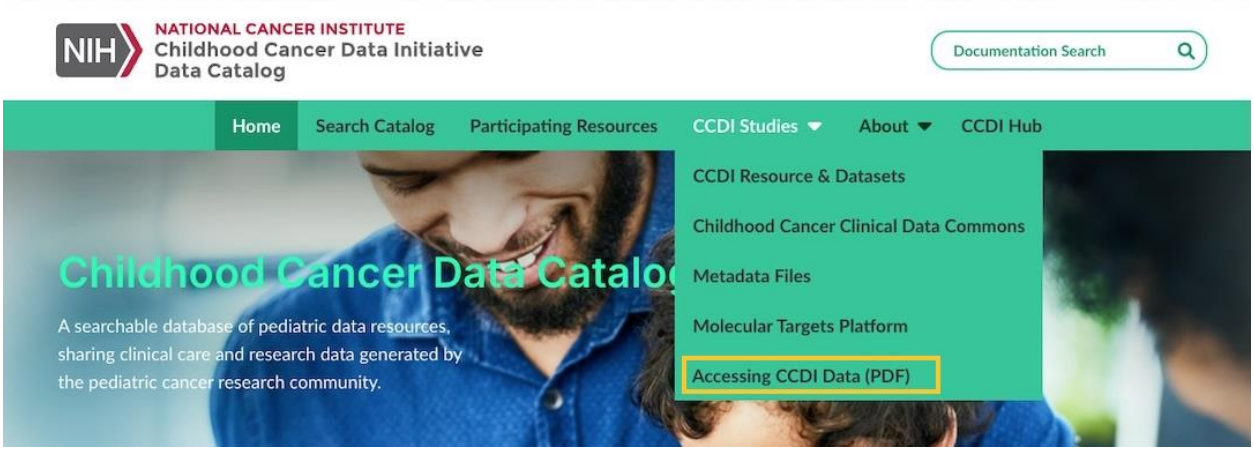

Figure 24: Childhood Cancer Data Initiative Page Selected for the CCDI Studies Link

About CCDI Data Catalog Page

The CCDC About CCDI Data Catalog page describes the CCDC and provides information on engaging with CCDI and contributing to the CCDC (Figure 25). The About page can be reached by navigating to the 'About CCDI Data Catalog' submenu item within 'About' dropdown on the menu bar. Once on the page, users can review information about the CCDC and how it fits within the CCDI. The About page also

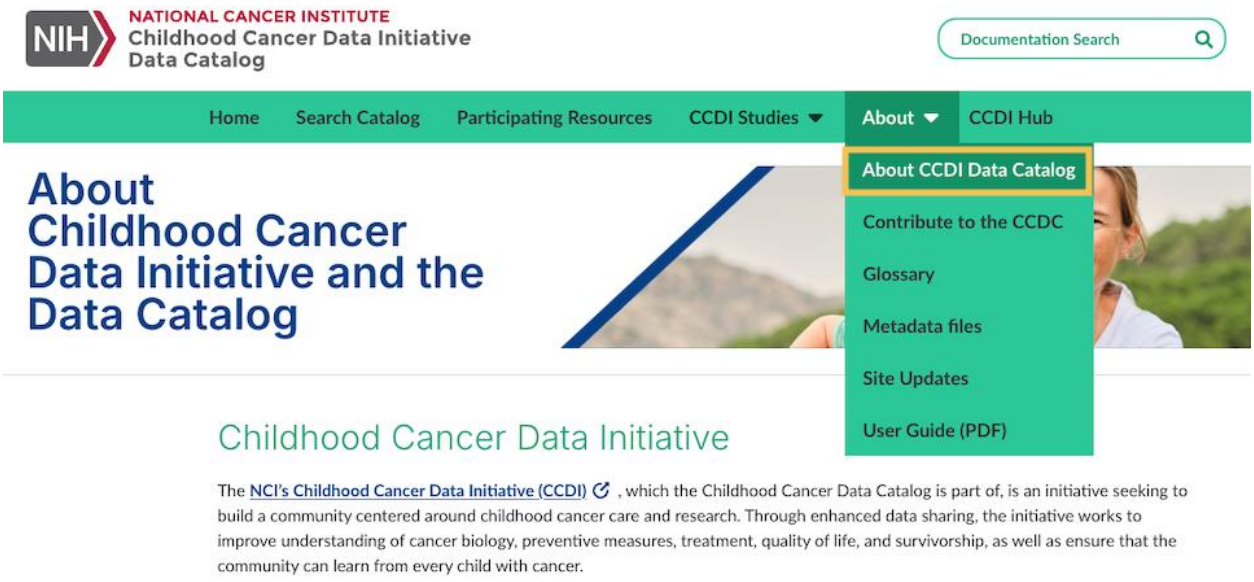

Figure 25: About CCDI Catalog Page

provides a link to join the CCDI mailing list. Joining the mailing list does not affect a user's experience on the CCDC site, but provides future news, content updates, and other information from the CCDI.

Contributing to the CCDI Data Catalog

The Contribute to the CCDI Data Catalog page provides information about submitting data and joining the CCDC's collection of community resources (Figure 26). The CCDC team is continually looking to expand and improve the inventory of pediatric oncology data resources on the CCDC site. Users with knowledge of additional resources are invited to share that knowledge with the CCDI team and contribute to the advancement of the CCDC. Users with comments are invited to contact the CCDI directly using the email link provided on the About page, using the Contact Us link in the site footer, or directly emailing NCIChildhoodCancerDataInitiative@mail.nih.gov.

Users willing to share a new resource are invited to download the summary submission template. This template is the tool used to add new information. Instructions on filling out this template are available by contacting the NCI's Office of Data Sharing (ODS) at NCIChildhoodCancerDataInitiative@mail.nih.gov. Completed templates can be submitted to the CCDI.

Please note, all submitted templates will be reviewed by the CCDC team. Any accepted data submission or suggested changes will be incorporated as part of periodic releases of the CCDC site.

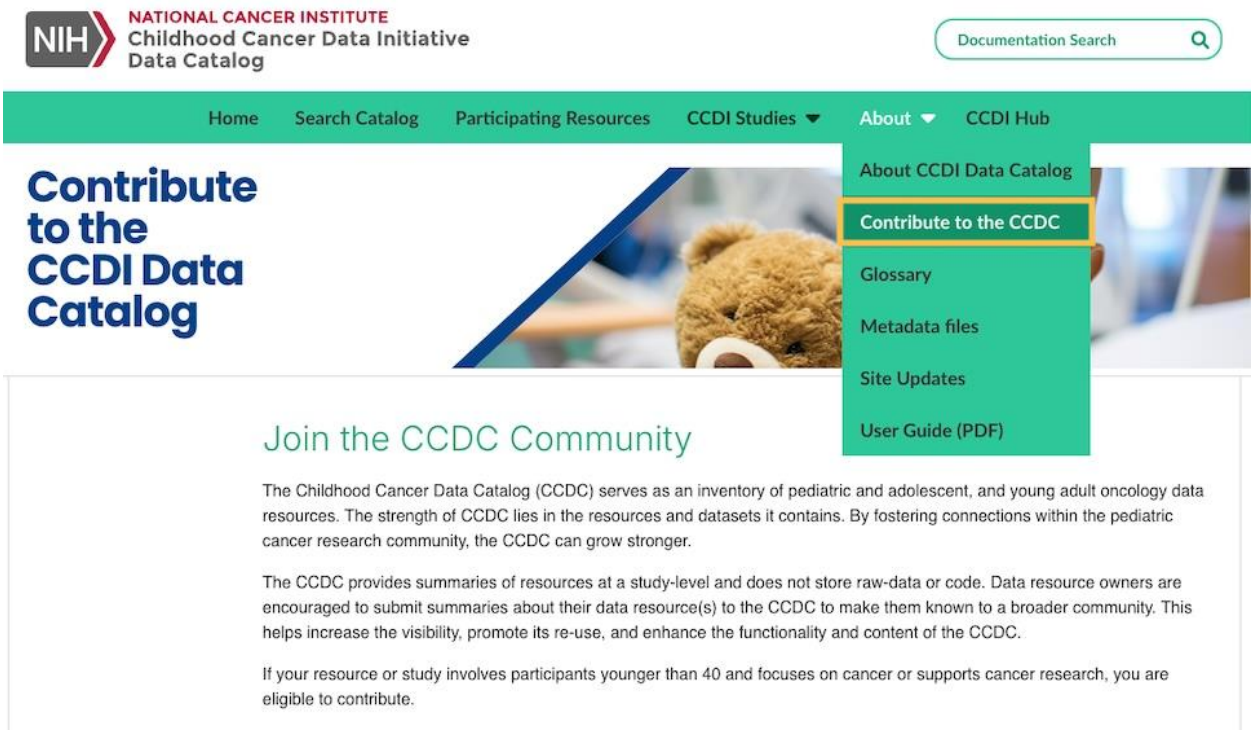

Figure 26: Contribute to the CCDI Data Catalog

Glossary Page

The CCDC Glossary page provides a listing of terms and abbreviations used on the CCDC site (Figure 27). To reach the Glossary page users navigate to the 'Glossary' submenu item within 'About' dropdown on the menu bar.

The glossary page has two main components, the alphabetical ribbon for navigation and the term pane where information is listed and defined. The information in the glossary is arranged alphabetically and can be navigated by selecting letters on the alphabetic ribbon. Letters on the ribbon will appear in one of three ways. If a letter appears green with an underline it indicates that the letter is selected. Terms from a selected letter will appear in the term pane. If a letter appears black, then that letter contains glossary terms but is not selected. Black letters can be selected to update the term pane with terms of the newly selected letter. If a letter appears grey, there are no terms defined for that letter.

Each glossary term will always have the following components:

- **Term** – The term being defined
- **Term Category** – A grouping of like terms by the CCDC
- **Definition** – The definition of the term

Terms may also have the following:

- **Reference** – The definition source if not defined by the CCDC team
- **Reference URL** – A URL link to the external definition

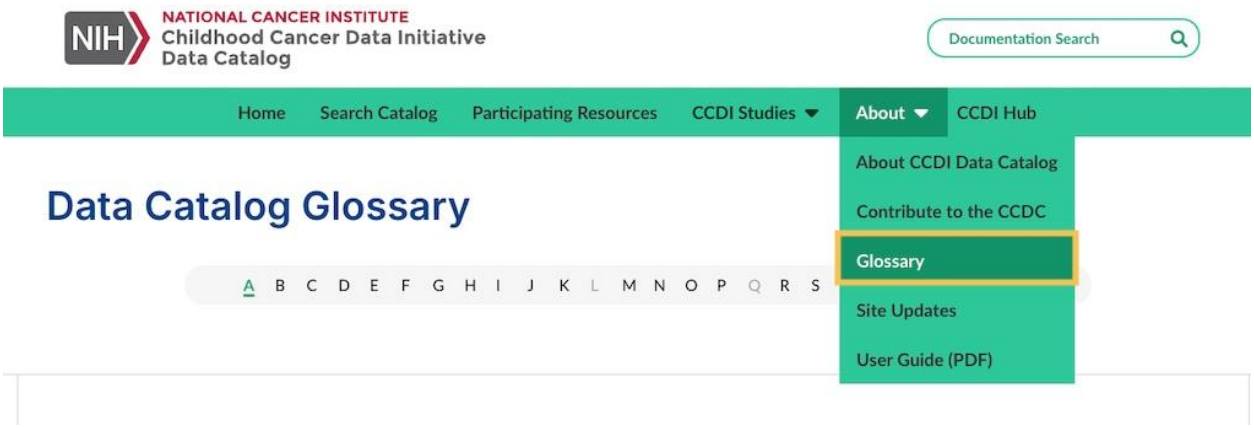

Figure 27: Data Catalog Glossary Page

Site Update Page

The CCDC Site Updates page provides a reverse chronological listing of updates made to the CCDC site (Figure 28). The Site Update page can be reached by either selecting an update on the Homepage *Featured Items* widget or by navigating to the 'Site Update' submenu item within 'About' dropdown on the menu bar. By default, the most recent update will display, but can be changed to display details about any release by selecting the corresponding item from the lefthand menu.

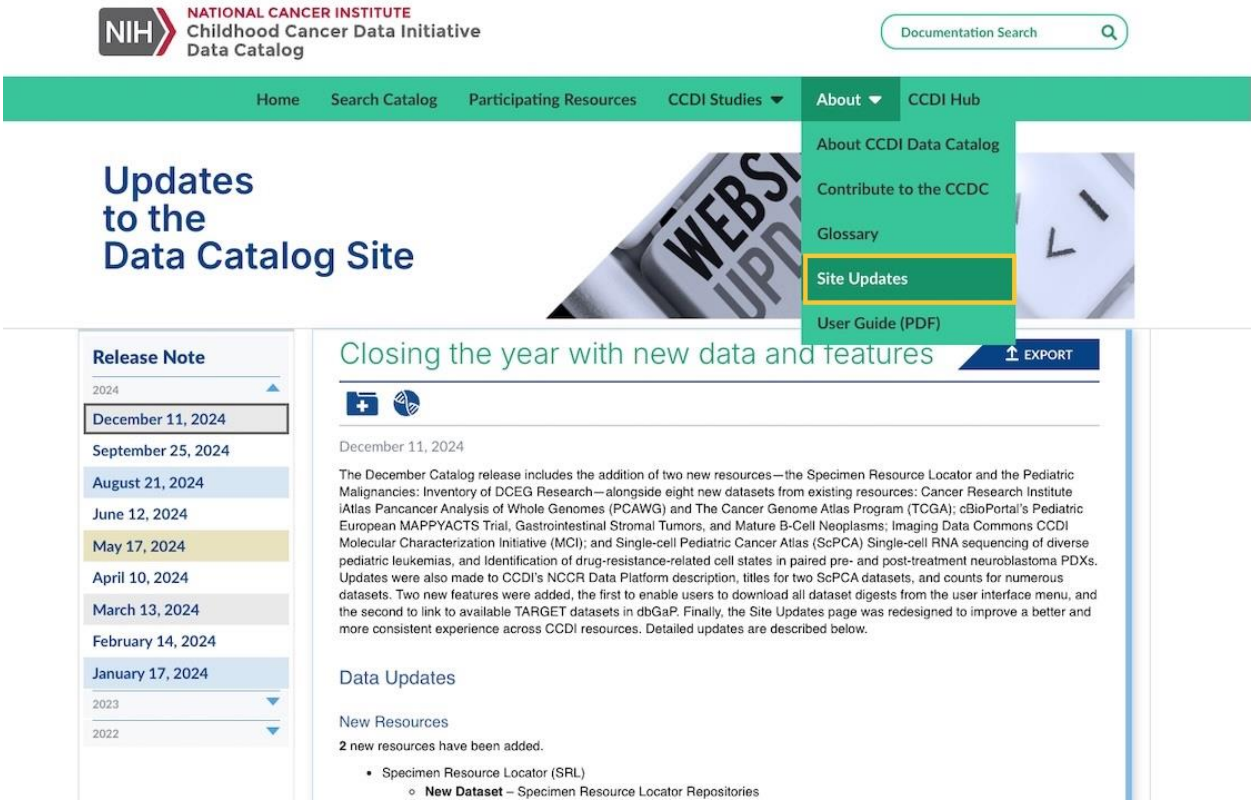

Figure 28: Data Catalog Site Update page

Each entry on the Site Update page shares information about changes to site data and site features for a specified release (Figure 29). At the top of each entry is an icon(s) to identify the Data Content Type included in the update. These new icons convey information across all the newly added datasets at a quick glance. Each update entry appears as its own scrollable listing on the site updates page. Entries can be navigated using the scroll bar on the right of each entry. Information within each entry is displayed with additional formatting to further clarify and break down the updates. Each site update can be exported as a PDF by selecting 'Export' in the upper right of each entry.

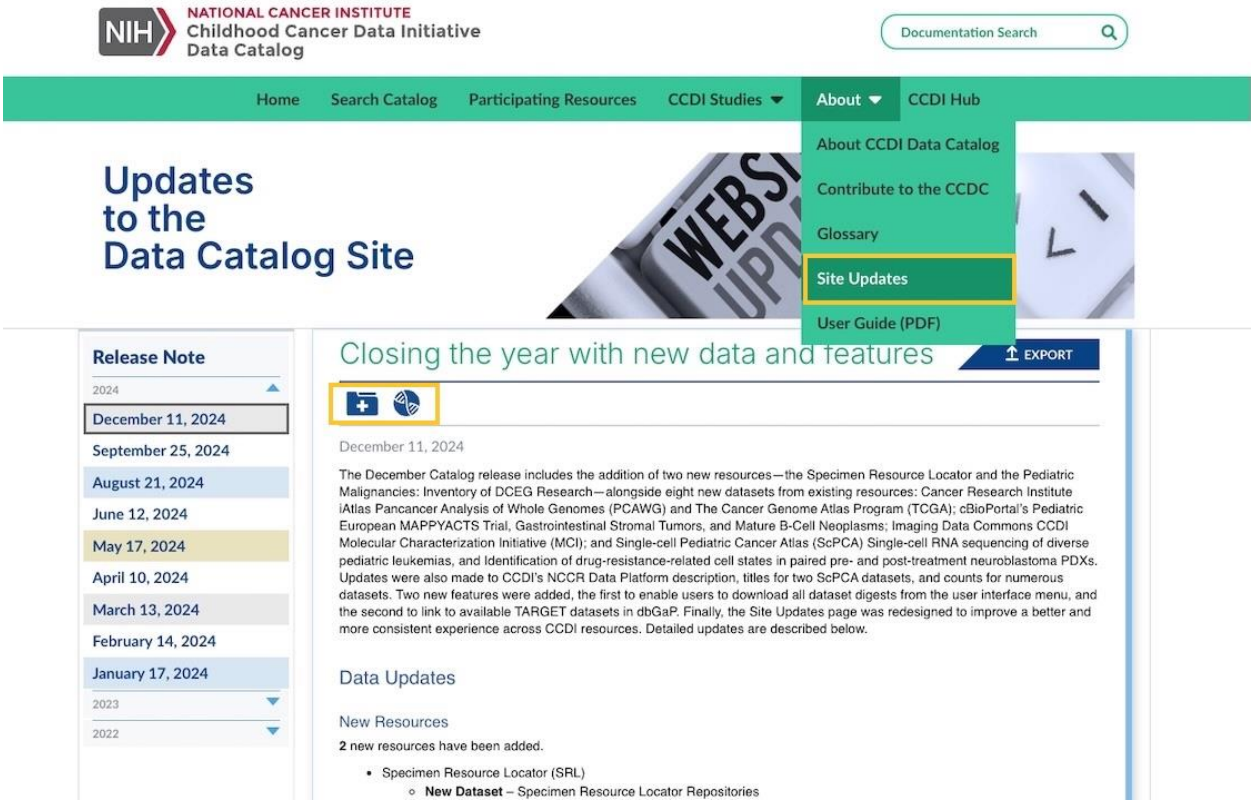

Figure 29: Data Catalog Site Update page – Data Content Type icon and right-side scrolling feature

User Guide

The CCDC User Guide link provides a link to a PDF version of the site user guide (Figure 30). Selecting the user guide link should open the guide in a new tab. Please note, for Safari user, clicking on the "User Guide" menu item will directly download the PDF format of the user guide.

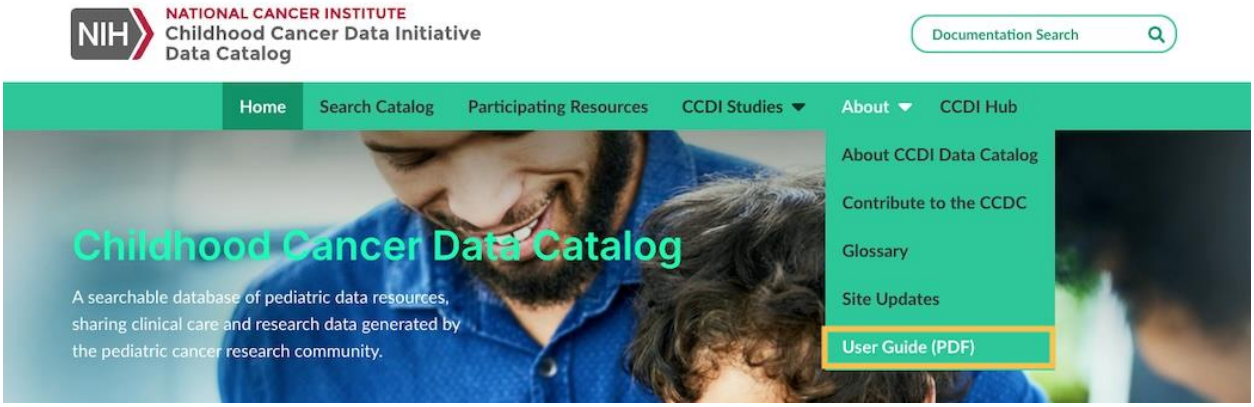

Figure 30: User Guide Page

CCDI Hub Page:

The Home Page Navigation bar provides a direct link to this CCDI Applications (Figure 31):

- **CCDI Hub**: The Childhood Cancer Data Initiative Hub (CCDI Hub) is an entry point to the Data Ecosystem and provides direct links to applications, tools, resources, and additional technical information, including the CCDC.

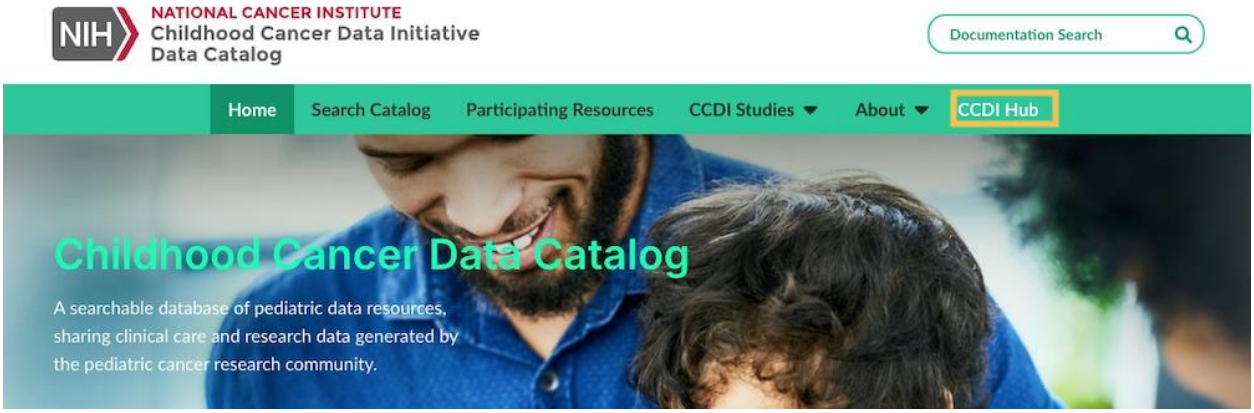

Figure 31: Link to CCDI Hub from Navigation bar

Documentation Search Page

The Documentation Search field is a free-text search tool that searches across three CCDC site pages – Home page, About CCDI Data Catalog, and the Glossary. The search, located on the right-hand side of the CCDC site header, is accessible on any CCDC site page (Figure 32). To use the documentation search, type in a term or text of interest and press enter or click on the magnifier icon. A successful execution

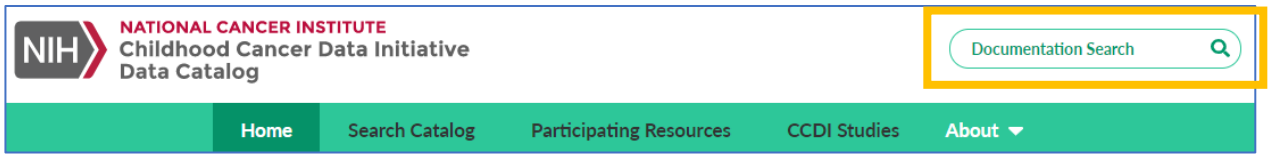

Figure 32: Documentation Search Field

will take you to the documentation search page. This page is only accessible when using the Documentation Search field.

Parameters for the Documentation Search:

- A minimum of 3 characters are needed for a term search.
- Searches will return both full and partial word results (i.e., “can” returns results for cancer).
- If you search multiple terms (i.e., “Cancer Institute”) the search result returns sources that contain all the specified terms (AND Boolean operator).

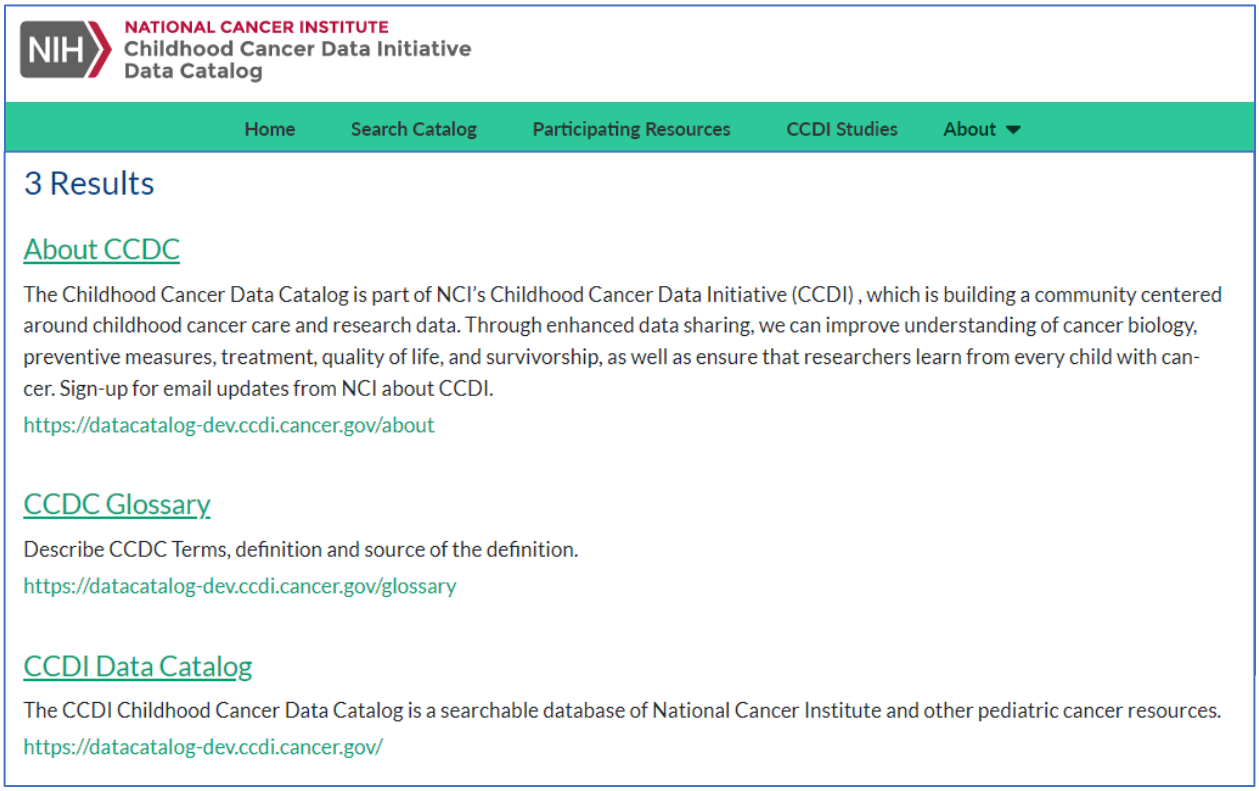

Figure 33: Documentation Search Page Results Pane

The result of a search appears in the results pane on the main body of the page (Figure 33). Results only return if a static page includes the searched term. For each returned entry, the search displays the page title, the first block of page text, and the URL of the page. To clear the search field, select the ‘X’ on the right side. Clearing the field will remove typed text but will not reset an executed search.

Additional Support

Users requiring any additional support or with any questions about the CCDC can contact Childhood Cancer Data Initiative at NCIChildhoodCancerDataInitiative@mail.nih.gov. Please include a name and a reference to the CCDC with submitted any inquiry.

Appendix A: Participating Resources

Participating Resource	Resource Type	Data Update Date
Cancer Genomics Cloud (CGC)	Analytic Tool	12/14/2022
Cancer Dependency Map (DepMap)	Analytic Tool	4/27/2023
Cancer Research Institute iAtlas (CRI iAtlas)	Analytic Tool	11/12/2024
HitWalker2	Analytic Tool	12/20/2022
Osteosarcoma Explorer (OSE)	Analytic Tool	11/30/2023
The Cancer Proteome Atlas (TCPA)	Analytic Tool	12/19/2024
Vizome	Analytic Tool	12/19/2022
Xena Browser	Analytic Tool	4/12/2024
NCTN Biobank	Biorepository	1/17/2025
Norwegian Childhood Cancer Biobank (NCCB)	Biorepository	8/3/2023
Swedish Childhood Tumor Biobank (BTB)	Biorepository	8/3/2023
The Biopathology Center (BPC)	Biorepository	8/3/2023
Specimen Resource Locator (SRL)	Biorepository	11/15/2024
VIVO Biobank	Biorepository	8/3/2023
Cancer Epidemiology Descriptive Cohort Database (CEDCD)	Catalog	7/10/2024
CancerModels.Org	Catalog	10/16/2024
Children's Oncology Group (COG)	Catalog	12/18/2023
Human Cancer Model Initiative (HCMI)	Catalog	9/6/2023
Inventory of Division of Cancer and Epidemiology and Genetics (DCEG) Research	Catalog	11/20/2024
MicroRNA Childhood Cancer Catalog (M3Cs)	Catalog	1/3/2023
National Clinical Trials Network (NCTN) Navigator	Catalog	11/27/2024
cBioPortal for Cancer Genomics	Data Repository	11/24/2024
Childhood Cancer Repository	Data Repository	6/17/2024
Genomics Evidence Neoplasia Information Exchange (GENIE)	Data Repository	10/11/2024
Gene Expression Omnibus (GEO)	Data Repository	5/28/2024
Genomic Data Commons (GDC)	Data Repository	10/01/2024
Imaging Data Commons (IDC)	Data Repository	10/10/2024
International Cancer Genome Consortium (ICGC)	Data Repository	7/1/2024
Kids First Data Resource	Data Repository	12/30/2024

Oncogenomics	Data Repository	2/14/2023
Patient-Derived Xenograft and Advanced In Vivo Models (PDX-AIM)	Data Repository	5/21/2024
PDX Development and Trial Centers Research Network (PDXNet)	Data Repository	4/12/2024
PDX for Childhood Cancer Therapeutics (PCAT)	Data Repository	12/19/2024
PedcBioPortal for Integrated Childhood Cancer Genomics (PEDC)	Data Repository	12/23/2022
Pediatric Cancer Data Commons (PCDC)	Data Repository	3/14/2024
St. Jude Cloud	Data Repository	12/31/2024
The Cancer Imaging Archive (TCIA)	Data Repository	7/17/2024
The database of Genotypes and Phenotypes (dbGaP)	Data Repository	12/15/2024
The Jackson Laboratory PDX Models (JAX PDX)	Data Repository	3/22/2023
Treehouse Childhood Cancer Initiative	Data Repository	3/22/2023
Victorian Paediatric Cancer Consortium (VPCC)	Data Repository	10/25/2023
Clinical Interpretations of Variants in Cancer (CIViC)	Knowledgebase	3/17/2022
Molecular Targets Platform (MTP)	Knowledgebase	8/4/2022
OncoKB	Knowledgebase	3/20/2024
Cancer Genome Characterization Initiative (CGCI)	Program	7/23/2024
Center for Pediatric Tumor Cell Atlas (HTAN)	Program	8/9/2024
Childhood Cancer Data Initiative (CCDI)	Program	11/21/2024
Childhood Cancer Survivor Study (CCSS)	Program	10/03/2022
Innovative Therapies for Children with Cancer Paediatric Preclinical Proof-of-Concept Platform (ITCC-P4)	Program	6/17/2024
My Pediatric and Adult Rare Tumor network (MyPART)	Program	8/25/2023
Pediatric MATCH	Program	7/25/2024
Pediatric Preclinical Testing Consortium (PPTC)	Program	8/5/2024
Single-cell Pediatric Cancer Atlas Portal (ScPCA)	Program	11/20/2024
Texas KidsCanSeq	Program	2/16/2023
Texas Pediatric Cancer Drug Testing Core & Pediatric Preclinical In Vivo Testing Consortium	Program	8/23/2023
Therapeutically Applicable Research to Generate Effective Treatments (TARGET)	Program	10/22/2024
WHO IARC	Program	5/3/2024
Individualized Cancer Therapy Study (iCat)	Project	7/20/2022
All of Us	Registry	2/15/2023
Central Brain Tumor Registry of the United States (CBTRUS)	Registry	3/20/2024
Fibrolamellar Registry	Registry	2/23/2024

National Childhood Cancer Registry Database in SEER*Stat (NCCR)	Registry	1/2/2024
Population-based Research to Optimize the Screening Process (PROSPR)	Registry	8/24/2023
South African National Cancer Registry	Registry	4/25/2024