

CHAPTER 3

SEARCHING A CAINTEGRATOR2 STUDY

This [chaptersection](#) describes the processes for searching studies within calIntegrator2 using the search and browse tools.

Topics in this [chaptersection](#) include:

- [Search Overview on this page](#)
- [Launching a Search](#) on page 28
- [Managing Queries](#) on page 33

Search Overview

The search and browse functions in calIntegrator2 allow you to search for clinical data, genomic or imaging data that were uploaded into the application as part of a study. When gene expression and imaging data are uploaded into a calIntegrator2 study, mapping files that correlate the data in those files to patient IDs in the clinical data file must also be uploaded. When you launch a search, calIntegrator2 finds and integrates the clinical, genomic and imaging data based on the mapping files and the criteria that you define in the search query.

In a search query, you can specify criteria for just one of the data types, or configure complex search criteria that join two or three data types. The available criteria for the query were defined when the study was deployed.

The basic workflow for a study search follows these steps:

1. Select the study to be searched.
2. For a simple search, select one data type:
 - **Clinical:** searches one or more uploaded CSV files for data identifiers or annotations (column headers) specified when the study was created

- **Genomic:** searches caArray experiments samples uploaded in the study for gene expression data by gene name or reporter ID.
 - **Image Series:** searches NCIA files uploaded in the study for image annotations or links to images, identified by subject identifiers or image series IDs
3. Define criteria for the search in the selected data type and run the search.
 4. For a more complex search, select multiple criteria from more than one data type.
 5. Specify whether you want clinical/imaging annotations to display or genomic data to display.
 6. Review search results.
 7. Configure results column and sorting display settings. You can do this before or after you run a search. If you choose to do it after, you must re-run the search.
 8. Download annotation search results as a CSV file. The CSV file contains only the data you specified in the annotation and display configurations.
 9. Follows links to NCIA in the search results to view or download images located in the search.

Launching a Search

Managing Queries

Launching a Search

To initiate a search of all annotations and/or other data in a study, follow these steps:

1. In caIntegrator2, in the upper right hand corner, select the study you want to browse or perform a simple search.
2. On the left sidebar, under the first section that displays the [study name], click **Search [Study Name]**. This opens a simple search query page with five tabs.

Search Rembrandt/VASARI

The screenshot shows a web interface for searching data. At the top, there are five tabs: 'Criteria' (selected), 'Columns', 'Sorting', 'Query Results', and 'Save as...'. Below the tabs is a section titled 'Define Query Criteria'. It contains a dropdown menu with 'Clinical' selected, followed by an 'Add' button with a plus icon. Below this, a message states 'No criteria added. Please select criteria from the pulldown box.' At the bottom of the criteria section, there are radio buttons for 'or' and 'and', with 'and' being selected. A 'Run Query' button is located at the bottom right of the interface.

3. On the Criteria tab, in the drop-down list, select the type of data you want to search. The listed options reflect the type of data that have been uploaded to the study.

Note: You can perform a search using one or more criteria you set in one of the data types, or you can define criteria in more than one data type per query, creating a more complex search.

- **Clinical**
- **Gene Expression**
- **Image Series**

Note: NCIA submissions are organized in the following hierarchy, which illustrates the relationship of an image series to its parent study and patient, as well as to the images in the series.:

Clinical trial > Patient (Subject) > Study > Series > Images

4. Click the **Add** button to define annotation elements for the search.

Clinical and Image Series

- If you select Clinical or Image Series data types, an additional drop-down list displays data elements that are annotation definitions specified when the data was uploaded into the study. Select a search criterion from among the options. You can make only one selection at a time.

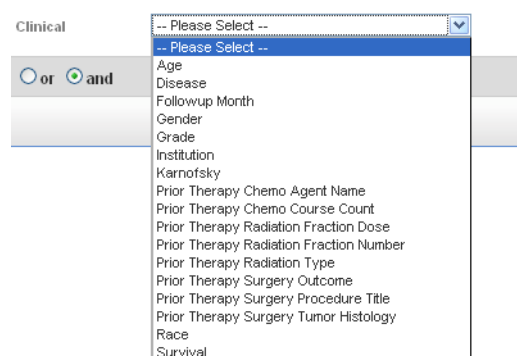


Figure 3.1 Additional clinical search criteria

- Each choice opens other fields relevant to the selection where you can further define your search query.
 - If permissible values were added when the annotation was defined, you can ? or must? select among the values in a drop-list that displays on the right side of the page.

- If no permissible values were defined as part of the annotation, you have the option to enter descriptive text in a text box on the right side of the page. ([X ref](#)).

The screenshot shows a web interface titled "Define Search Criteria". At the top, there is a dropdown menu set to "Clinical" and an "Add" button with a plus icon. Below this, there are two rows of criteria. The first row has "Clinical" in a dropdown, "Disease" in a dropdown, and "equals" in a dropdown. The second row has "or" selected in a radio button group (with "and" also available) and a list of values: "ASTROCYTOMA", "ASTROCYTOMA", "CELL_LINE", "GBM", "OLIGODENDROGLIOMA", and "UNCLASSIFIED". At the bottom right, there is a "Run Search" button.

Figure 3.2 You may be able to further define search criteria when you select a specific clinical or imaging annotation element

Note: When working with image data, if only an Imaging Mapping file was uploaded when the study was created and not an Image Series Annotation file, you cannot enter image search criteria. The search results will, however, display a link that allows you to view the associated images in NCIA.

Gene Expression

Searches for gene expressions data will return results which you are permitted to see.

- For the Gene Expression selection, you can select **Gene Name** or **Fold Change**. Fold change is used to identify genes with expression differences compared to control samples, as defined when the study was deployed in calIntegrator2. You can enter query values in greater/lesser-than-or-equal-to arguments.
- 5. To add additional criteria for the search, repeat steps 3 and 4, as appropriate. You can set more than one data type or more than one criterion for a data type. The criteria become cumulative, thus refining the search.
- 6. Once you have configured the query criteria, select the Boolean **Or** or **And** search operator at the bottom of the page.
 - **Or** finds a data subset with at least one of the search criteria
 - **And** finds a data subset with both/or all search criteria.
- 7. Click the **Remove** button to clear any data elements you have defined.
- 8. You can launch the search from this tab. Click the **Run Search** button. For information about the search results, see [Chapter 4 Viewing Search Results](#). You may want to run the search first to see what kind of results you get before you configure the data display, described in step 9.

– or –

- 9. On the Columns tab, you can specify the columns you want to display in the search results data. On the Sorting tab, you can specify how the data is to be sorted. By default, only the column showing the [subject] identifier displays. For more information, see [Columns and Sorting Tabs](#) on page 31.

Note: As long as you are still in the current query session, you can return to the Criteria, Columns and Sorting tab to add, modify or remove data and display criteria and re-run the search. If you save the query, your current search criteria are saved.

Saving a Query

Browsing Search Results

Exporting Data

Columns and Sorting Tabs

You can specify columns and sorting options for the way you want the search results to display either before you run the search, or after you run the search. If you run the search directly from the Criteria tab before setting the columns/sorting features, by default only the Subject Identifiers display on the Search Results tab. You can then come back to the Columns and Sorting tabs to expand the display options and re-run the search, having set the display parameters.

The selection you make on the Columns tab determines whether calIntegrator2 displays search results for clinical or genomic data. It filters the search based on the criteria you set on the Criteria tab, whether it is clinical, gene expression or image series data type(s). In other words, if you select clinical criteria on the Criteria tab, but select Genomic on the Columns tab, the data subset that displays on the Search Results tab is genomic data that is filtered by the clinical criteria you defined on the Criteria tab.

1. On the Columns tab, select the **Clinical** or **Genomic** radio button to search clinical data.

Search Rembrandt with NCRI Study

Clinical – Select the annotation elements that you want to display in the search results. All elements listed are column headers in the data uploaded to the study. You can make multiple selections on this list.

Note: For Clinical Annotations, the Patient or Subject Identifier display by default in the search results.

Results display as tabular data.

Genomic – Select the Reporter Type define:

- Gene Name--finds all reporter that match criteria for gene you identified on the Criteria tab
- Reporter ID--finds all reporters that match criteria identified on Criteria tab

Results display in a gene expression data matrix.

Imaging – If imaging annotations have been added to the study, annotation elements also display on the lower right section of this page when you select **Clinical**. All elements listed are column headers in the image annotation data uploaded to the study. You can make multiple selections on this list.

Note: If you select even one Image Annotation, the Image Series IDs display by default in the search results. If you select no Image Annotations, however, even if you have selected image series criteria on the Criteria tab, no image series IDs display in the search results. The fact that images can be located, however, in NCIA is indicated by two image-related buttons at the bottom of the Query Results page. You can open the images in NCIA, but they will be at StudyInstance UID level.

Results display as tabular data.

2. Use the **Select All** or **Unselect All** buttons to aid you in making your selections.

The column selection is saved as part of the query if you save it. See [Saving a Query](#) on page 33.

3. Select the Sorting tab and indicate the column order of the Search Results.

Search Rembrandt with NCRI Study

Criteria	Columns	Sorting	Query Results	Save as...
Set Sort Order for Selected Columns				
Column	Order (L-R)			
Tumor Location	1			
GENDER	2			
Survival	3			
Age	4			
Run Query				

Sorting parameters are saved as part of the query if you choose to save it using the Save Query feature. On the Search Results page, you can also sort the results by clicking on a column name.

4. Click **Run Search**. Search results display on the Search Results page. For information about the search results, see [Chapter 4 Viewing Search Results](#).

Launching a Search

Saving a Query

Editing a Query

Browsing Search Results

Managing Queries

When you create a search query in calIntegrator2, you can save the query for later use or edit it. [For more information, see these topics:](#)

[Saving a Query](#)

[Editing a Query](#)

Saving a Query

1. Click the **Save As** tab and enter a **Search Name** and **Search Description**, unique to the search. *Example: **Batch ID 6 and female***
2. Check the appropriate options to save data type criteria as well as the column and sort selections you have defined for the search.
3. Click **Save**.

Once the query is saved, it is listed by its name under the **Study Data > Queries > My Queries** in the left sidebar, whenever the study to which the query applies is selected in calIntegrator2.

[Launching a Search](#)

[Columns and Sorting Tabs](#)

[Editing a Query](#)

[Browsing Search Results](#)

Editing a Query

1. To edit a query, select it in the left sidebar under the **Study Data > Queries > My Queries**.
2. Change the query and display criteria on the Criteria, Columns and Sorting tabs.
3. On the Save As tab, check the appropriate options and click **Save As**. You can use the same name as the original query or modify the name as needed.

[Launching a Search](#)

[Columns and Sorting Tabs](#)

[Saving a Query](#)

[Browsing Search Results](#)

Exporting Query Results

After running a search, you can export the result set or a subset as a tab-delimited text file. For more information, see [Exporting Data](#) on page 42.

[Launching a Search](#)

[Columns and Sorting Tabs](#)

[Saving a Query](#)

[Browsing Search Results](#)

