

## Chapter 20: Data Searches and Queries

There are three search tools in SEER\*DMS: Quick Search, Patient Lookup, and Data Search. Use the Quick Search or Patient Lookup when you are looking for a specific patient's data. Use the Data Search to review a set of Patient Sets or records; to create an extract file containing Patient Set or record data fields, or to execute SQL queries against the SEER\*DMS database.

- **Quick Search** – If your search is based on a single data field, enter the field's value into the Lookup box on the user bar. You may search for a patient's data by entering a Patient Set ID, Record ID, Import ID, patient name, date of birth, or SSN. The Quick Search also provides a shortcut for opening an AFL, follow-back need or group, contact, facility, or worklist task. It is a shortcut to the Data Search for searches based on Import ID or a list of Patient Set or Record IDs.
- **Patient Lookup** – Use the Patient Lookup to search for a patient's data based on multiple search fields or to search on fields not supported by the Quick Search. You may enter search values for standard patient identifiers such as name, SSN, date of birth, sex, accession number, and medical record number.
- **Data Search** – The Record and Patient Set Data Searches allow you to define robust search criteria based on any record or Patient Set field, Import ID, Special Study ID, or type of worklist task. You may search for a set of records or Patient Sets by entering a list of Patient Set/Record IDs. The SQL Data Search can be used to run read only queries involving any table or view in the SEER\*DMS database. Use any of the Data Searches to assign Patient Sets or records to Quality Control tasks, to review unusual data, or to create small ad hoc extract files.

Once you find the data of interest, you may view or edit the data in the record or Patient Set editor. However, you must take precautions to ensure that you do not modify data that are being processed by others. To avoid the loss of effort, you should verify that the data are not involved in worklist tasks and are not being edited by another staff member (please refer to the *Direct Editing of Records and Patient Sets* section of this chapter for more information).

In this chapter, you'll learn about

- Quick Search
- Patient Lookup
- Data Search
  - Accessing the Data Search
  - Creating and Using Data Searches
  - Patient Set Data Search
  - Record Data Search
  - SQL Data Searches
  - Reviewing Results in the Data Search
  - Exporting Data Search Results
  - Running or Changing Data Searches
  - Saving and Closing Data Searches
  - Creating Workflow Tasks via the Data Search
  - SQL Tools
  - Database Documentation
  - Data Search Instructions and Help
- Direct Editing of Records and Patient Sets

## Quick Search

The Lookup search box on the SEER\*DMS user bar allows you to do quick searches. It is only available if you have system permissions to view and/or edit patient data.



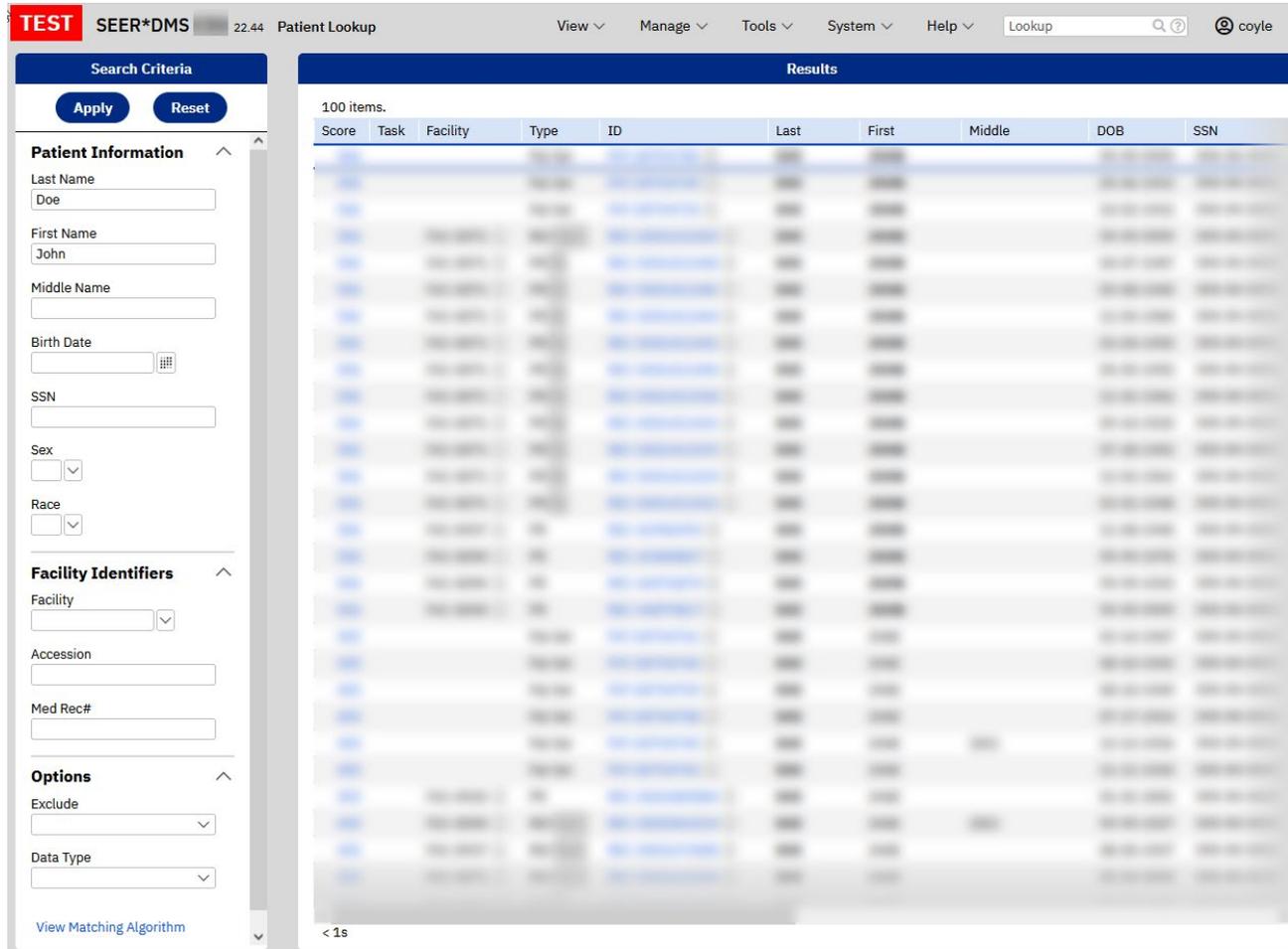
Enter the search text into the Lookup box and press Enter. If you enter a string consisting of all digits, the search will assume that you are searching by Record or Patient Set ID. You can view the complete list of Lookup search fields by clicking the question mark in the search box.

Commonly used search fields are listed below.

- **Patient Set or Record ID** – To open a Patient Set or record in the editor, simply enter the ID into the Lookup box. The ID prefix is optional (PAT-12345678 or 12345678 will work). The Patient Set or record matching the ID will open automatically. To search for a set of Patient Sets or records, enter a list of IDs separated by blanks or commas. The Data Search will open and the ID filter set. The Data Search does not allow you to search for Patient Sets and records at the same time, therefore, the first ID will determine whether it is a record or Patient Set search.
- **Patient Name** – You may enter “last name” or “last name, first name” into the Lookup box. The Patient Lookup will open and display the results of your search.
- **Date of Birth** – MM-DD-YYYY format is required. A valid month and year are required by the Quick Search. You may enter 99 for unknown day. The Patient Lookup will open with the results of your search. Patient Lookup results are limited to 100 matches.
- **Social Security Number** – A complete SSN (NNN-NN-NNNN, dashes required).
- **Import ID** – Enter an Import ID into the Lookup box to search for records in that import. The IMP prefix is required. The Data Search will automatically open. The fields displayed by default are determined by registry configuration settings. Data Search permissions are required for searches by Import ID.
- **Worklist Task ID** – Enter a single Worklist Task ID with the TSK prefix. The worklist will open and the worklist filters will be set to show the specified task.
- **AFL, Contact, Facility, or Follow-back Need ID** – If you have the appropriate permissions, you may open an AFL, Contact, Facility, or Follow-back Need by entering the appropriate ID. The AFL, PER, FAC, or FBN prefix is required. This allows you to bypass the manager when you wish to view a specific AFL, Contact, Facility or Follow-back Need.
- **AFL or Follow-back Group ID** – Enter a single Follow-back or AFL Group ID (the FBG or AFL prefix is required). The AFL or Follow-back Manager will automatically open and the filters will be set to show the items in that group.
- **Electronic Physician Report ID** – Electronic physician reports in the Clinical Document Architecture (CDA) format are stored in SEER\*DMS with an EHR prefix.
- **Novel Data Types** – SEER\*DMS supports data from novel data sources that can be linked to the standard cancer registry data. These currently include insurance claims data (CLM), pharmacy fill data (PHM), radiation oncology data (RAD), and genomic test results (GNM).

# Patient Lookup

The Patient Lookup can be accessed from the View menu or by pressing Enter in the Lookup search box at the top of the screen. The Patient Lookup enables you to search for a patient's data based on personal identifiers (Name, DOB, SSN, Sex, Race, Accession Number, Medical Record Number). The Data Type, Facility, and Exclude criteria can be used to limit the search to a subset of data.



## Search Fields in the Patient Lookup

SEER\*DMS uses registry-specific matching algorithms to compare the patient information specified in the search fields to Patient Sets and unlinked records in the database. The Patient Lookup matching algorithm is documented on the Matching help page in SEER\*DMS.

A score is calculated for each possible match. The score allows SEER\*DMS to sort the results by the likelihood of the match. Entering multiple identifiers quickens the search and increases the score of true matches. It is recommended that you enter information into as many fields as possible or enter a unique identifier, such as Social Security Number. If searching for a patient with a common name, it is particularly important to enter values into as many search fields as are known. At a minimum, you must enter one of these fields: SSN, Last Name, DOB, Accession Number, or Medical Record Number.

The Patient Lookup search fields are listed below.

- **Name (L,F,M)** – Three search fields are provided for searching by patient name. You may enter a full or partial name in any or all name fields. The fields are displayed in the order of last name, first name, middle name. However, you may enter “Last, First” into the last name field. SEER\*DMS will automatically parse the text into the two fields for the search.
- **Birth Date** – Enter date of birth in MMDDYYYY or MM-DD-YYYY format. If you only know the patient’s year of birth, enter 99 for day and month (e.g., 99-99-1961). A valid month and year are required when other patient identifiers are not specified (last name or social security number).
- **SSN** – Only complete Social Security Numbers may be entered. The value may be formatted as NNNNNNNNNN or NNN-NN-NNNN.
- **Sex** – A lookup enables you to search for data based on a valid code for sex.
- **Race** – A lookup enables you to search for data based on a valid code for race. One value for race may be selected.
- **Accession** – Only complete Accession Numbers may be entered. Accession Number is the facility’s identifier for a CTC. A search by Facility ID and Accession Number should return a single result. However, if you do not know the Facility ID or believe the facility may be coded incorrectly then you should search only by Accession Number.
- **Med Rec #** – Only complete Medical Record Numbers may be entered. Medical Record Number is the facility’s identifier for a patient. Two or more patients may have the same Medical Record Number if those MRNs were assigned by different facilities. You have the option of searching by Facility ID and MRN (one result should be returned); or you may search by MRN without the facility, but multiple results may be returned.

You must specify a value for at least one of these fields: last name, Social Security Number, Date of Birth, Accession Number, or Medical Record Number. The Patient Lookup is designed to search for data for a specific patient; you cannot use it to generate a generic list of data by sex, race, record type, or facility. To create such a list, you should either use the Data Search, as described later in this chapter, or query the database using an external tool.

### Limiting the Search

You may specify the records and Patient Sets to be searched by entering values for facility, data type, and the “exclude” field. Unlike the Search Fields, these criteria are never weighted. They are used to make a definite determination of whether a record or Patient Set is “in” or “out” of the search results.

- **Facility** – Limit the search to Patient Sets with admissions or treatments related to a specific facility (this may be the reporting or treating facility), or to unlinked records that were provided by the facility. Leave this field blank to search all data regardless of facility.
- **Exclude** – Exclude Patient Sets and records for patients who have been dead for more than the specified number of years (based on vital status and date of last contact). Leave this field blank to search all data regardless of vital status. The default setting for this field is determined by a User Preference setting. To change the default, click the Account link in the user toolbar (next to the Lookup search box).
- **Data Type** – Search all Patient Sets and unlinked records or limit the search to one or more data types. To include linked records in the search, you must include Patient Sets. The Patient Set to which the record is linked will be returned. Selecting a specific data type will return unlinked records.

## Executing the Search in the Patient Lookup

Enter as much information as known and click Search. Up to 100 matches will be displayed sorted by score with the best matches listed first. The results will include unlinked records and Patient Sets which yielded a score greater than zero when matched using the search and exclusion criteria. View the algorithm by clicking the *View Matching Algorithm* link shown below the search criteria.

You may click an ID to open a record or Patient Set in an editor. If the data are involved in a worklist task, a link will be displayed in the Tasks column. The link indicates the user who is assigned to the task or that the task is unassigned. Other data columns are described below.

To search for a new patient, click the Reset button. Patient information that you entered will be removed, the Facility field will be cleared, and the Exclude field will be set to the system default.

## Patient Lookup Search Results

The columns listed below are shown in the search results. The value in the table will be shown in bold if it is an exact match to a search field. If the matching algorithm supports partial matches for Social Security Number and Date of Birth, only the matching fragment will be shown in bold. Fields not listed may have been involved in the search. For example, the matching algorithm may compare last name to maiden name or alias last name.

- **Score** – A value calculated by the matching algorithm that allows the results to be sorted by the likelihood of the match. You may click the score to view a summary of the calculation. Results with the highest score are displayed first.
- **Task** – If the data are in a worklist task, a link will be displayed in the Tasks column. The link will be “unassigned” or the username of the person to whom it is assigned.
- **Facility** – If the data type is a record, then the reporting facility will be displayed.
- **Type** – Data type. Patient Set or a record type will be listed. If the match is a record and the record’s reportability status is not “reportable”, an icon will be displayed:
  - N Non-reportable
  - A Auditable
  - U Unknown (reportability flag not set; screening has not been completed)
  - N/A Not applicable; the record type is not screened (e.g., supplemental records)
  - No icon is displayed if the record is reportable.
- **ID** – Patient Set or Record ID. Click the Information Icon  next to the Record’s or Patient Set’s ID to see a select set of fields. To browse all pages of the record or Patient Set, click the Record’s or Patient Set’s ID.
- **Last Name, First Name, Middle** – Patient’s last, first, and middle name. If the text is an exact match to search criteria, it will be shown in bold. If the name is too long to display, an ellipsis will be shown at the end of the name (...).
- **DOB** – Date of birth.
- **SSN** – Social Security Number.
- **Sex** – Coded value for sex. Hold your mouse over the value to view text description.
- **Race** – Race indicated in the first race field of the Patient Set or record. Hold your mouse over the value to view text description.
- **DOLC** – Date of last contact.
- **VS** – Vital status.
- **MRN** – Medical record number.
- **Acc** – Accession number.
- **Address** – The patient’s current address.

## Accessing the Data Search

The Data Search can be accessed from the View menu. Use the Data Search if you are:

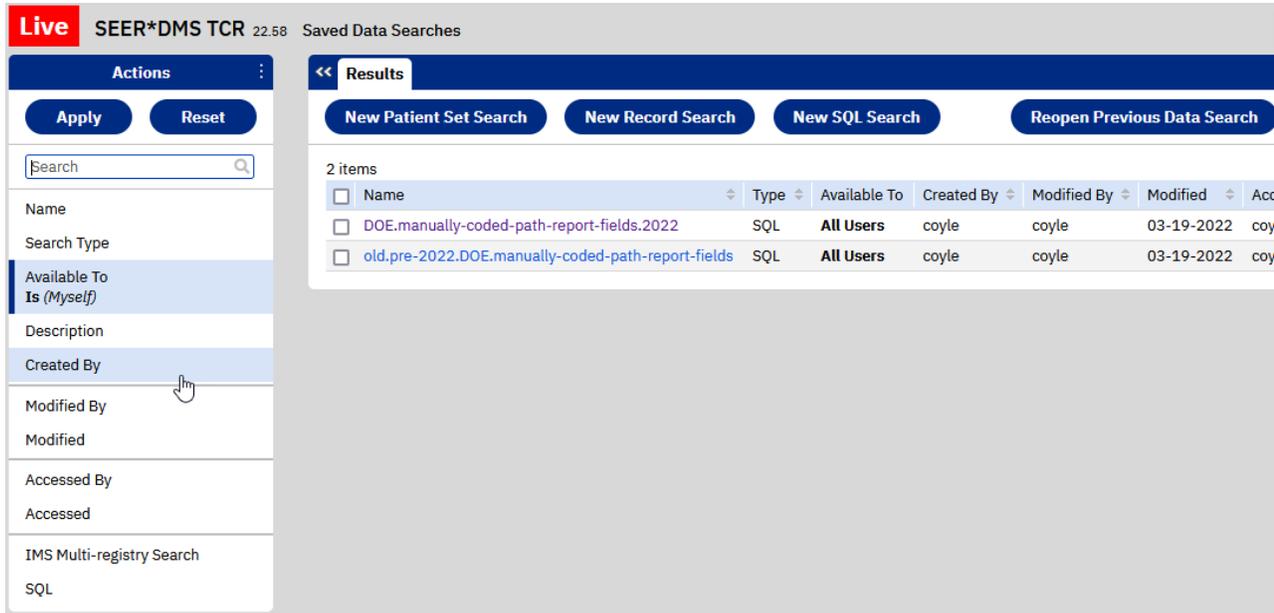
- looking for a set of records or Patient Sets that are similar in some way
- and/or you want to create a CSV extract of data fields
- and/or you want to create QC Tasks or Resolve Patient Set Errors tasks for a set of data

There are several system permissions that control access to the data search. Typically, coders should be given access to Patient Set and Record data searches, but system permissions related to exports and SQL searches should be assigned to a limited number of staff. The system permissions related to data searches are described below.

- *data\_search\_export* – Create export files from listings created in the Data Search. This permission is required to extract data from Patient Set, Record, or SQL searches.
- *data\_search\_filter\_manager* – Modify, delete, and assign saved searches (or filters) in the Data Search. This gives the user access to all saved searches that they can execute.
- *data\_search\_run\_pat* – Execute Patient Set Data Searches. A user with this permission can run Patient Set searches that are available to their account. They cannot create or modify Patient Set searches unless they also have the *data\_search\_write\_pat* permission.
- *data\_search\_run\_rec* – Execute Record Data Searches. A user with this permission can run Record searches that are available to their account. They cannot create or modify Record searches unless they also have the *data\_search\_write\_rec* permission.
- *data\_search\_run\_sql* – Execute SQL Data Searches. A user with this permission can run SQL searches available to their account. They cannot create SQL unless they also have the *data\_search\_write\_sql* permission. They cannot access a saved search unless someone assigns it to their account; or they have the *data\_search\_filter\_manager* permission.
- *data\_search\_write\_pat* – Create and execute Patient Set Data Searches. This permission allows the user to write data searches, and execute data searches assigned to their account, data searches assigned to their role, and data searches available to all users.
- *data\_search\_write\_rec* – Create and execute Record Data Searches. This permission allows the user to write data searches, and execute data searches assigned to their account, data searches assigned to their role, and data searches available to all users.
- *data\_search\_write\_sql* – Create SQL Data Searches. Execute SQL Data Searches assigned to the user's account, the user's role, or assigned to all users. A user with this permission can open another user's SQL data search, but cannot over-write it (the permission *data\_search\_filter\_manager* is required to modify someone else's data search). This permission provides access to all data (patient data, staff productivity, etc) and should be restricted to registry managers, QC managers, and staff who write SQL for managers.

# Creating and Using Data Searches

Saved Data Searches are displayed in a manager page as shown below. Use the fields in the left panel to find a saved Data Search. Use the buttons above the table to create a new data search. The option to Reopen Previous Data Search will only be shown when you had opened a data search in the same session (it will not remember the next time you login).



Use the left panel to find a saved search using fields associated with a saved Data Search:

- **Name** – the name of the saved data search.
- **Search Type** – indicates whether it is a Patient Set, Record, or SQL data search.
- **Available To** – this determines whether the data search can be accessed by all users, one specific user, or a group of users with the same role. Users with the data search filter manager permission can access searches assigned to other users. This is allowed so that user can provide support by editing or writing data searches for other users.
- **Description** – a brief description of the data search.
- **Created By** – username of the person who wrote the original version of the data search.
- **Modified By** – username of the last person to modify the data search.
- **Date Modified** – the latest save date for the data search.
- **Accessed By** – the username of the last person to open the data search.
- **Date Accessed** – the latest date that the data search was opened.
- **IMS Multi-registry Search** – this flag is set to TRUE if the data search was written by IMS staff and pushed to all registries. A multi-registry data search can only be modified by IMS staff; these are read only to registry staff in order to maintain consistency across registries.
- **SQL** – enter search terms to search for a data search based on the SQL text.

Use the buttons at the top of the main Data Search page to create a new Patient Set, Record, or SQL data search. This will determine the type of fields available in the left panel when you build a

search. Use the Patient Set Search for simple searches of consolidated data. Use the Record Search for simple searches of source record data. Create an SQL query to write a more complicated search, to specify criteria in consolidated and source data in the same search, or to write queries that use other tables in the database.

## Patient Set Data Search

This screen shot shows the controls for the **Patient Set Search**.

The screenshot shows a user interface for the Patient Set Search. At the top is a dark blue bar with the word "Actions" and a vertical ellipsis icon. Below this are three buttons: "Save", "Save As", and "Close". The main area is divided into sections: "Patient Set Search" with a "Name" field showing "not saved"; "Filters" with "Worklist Task Type", "Patient Set ID", and "Special Study" fields; a highlighted blue section for "Distinct Results" with "Is Yes" and a list of "Display Fields" including "displayId", "nameLast", "nameFirst", "nameMiddle", "birthDateMm", and "9 more..."; and an "Order By" field at the bottom.

**Name** is set when the search is saved.

**Worklist Task Type** can be used to search for records in a worklist task or exclude records in a worklist task.

Use the **Patient Set ID** control to search for a single ID or paste a long list of IDs into the filter. Each ID in a list must be separated by a blank or a comma. The PAT- prefix is optional.

Use **Special Study** to search for Patient Sets assigned to a specific special study.

Set **Distinct Results** to Yes to suppress duplicate rows in your results.

The **Display Fields** control defines the columns shown in the results table.

**Order By** determines the sort order of the table. Sorting results within SEER\*DMS may increase the time to display large result sets. If you are exporting data then consider leaving this blank and sorting in the external software.

## Record Data Search

This screen shot shows the controls for the **Record Search**.

**Actions**

Save Save As Close

**Record Search**

**Name**  
not saved

**Filters**

Worklist Task Type

Record ID

Import

Special Study

**Distinct Results**  
Is No

**Display Fields**  
documentId  
reportability  
dateOfDiagnosisYear  
primarySite  
laterality  
9 more...

Order By

**Name** is set when the search is saved.

**Worklist Task Type** can be used to search for records in a worklist task or exclude records in a worklist task.

The **Record ID** control works the same as other ID filters. You may enter a single ID or paste a long list of IDs into the filter. Each ID in a list must be separated by a blank or a comma. The PAT- prefix is optional.

Use the **Import** filter to select records from a specific import.

Use **Special Study** to search for Patient Sets assigned to a specific special study.

Set **Distinct Results** to Yes to suppress duplicate rows in your results.

**Display Fields** defines the columns shown in the results table.

**Order By** sets the sort order of the table.

### Distinct Results

Distinct Results will prevent you from having duplicates in your results. *It comes into play when the search criteria fields are at a more detailed level than the display fields.* It is recommended that you set it to Yes in all Patient Set Searches, but setting to No will speed up some complicated searches.

Recommendation: Set Distinct Results to yes. Only set it to no if the query is very slow and you either know that your search cannot include duplicate results; or you do not mind having duplicates in your results.

Before setting it to No, be sure that you understand that you may have duplicate rows in your results. Review the example below. The criteria include a CTC-level field (year of diagnosis) but the display fields are all at the patient level. As an example, John Doe and the Patient Set ID will be listed twice if he had a prostate and a lung cancer diagnosis during those years. If you do want a separate result for each cancer then you should include cancer fields in the result fields (you

could include tumor record number, sequence number, or site). If you only want patient name and ID, set distinct to yes so that you only get one result per patient.

**Example:**

- You search for CTCs dx'd between 1973-2008
- And you set your display fields to "Patient Set ID" and "Name"
- Summary: search criteria include a CTC-level field, but display fields are all patient level.

If Distinct Results = No, you will get duplicates for patients who have 2 or more diagnoses in 1973-2008. "PAT-12345678, John Doe" could be listed twice.

**Criteria**

Criteria determine which Patient Sets or records are returned in the results. You can create a fairly complex multi-line statement. The lines are joined with conjunctions: "not", "and", "and not", "or", "or not"; and you can enter parentheses when the Boolean logic requires it. When you are finished defining your criteria, click **Run** (or press Ctrl + Enter) to run your search.

Your search will return a result set based on the full criteria statement "and-ed" with all other filters. For example, the result set for a Patient Set Search is determined as follows.

Patient Set Results = *Criteria Statement + Worklist Task Filter + ID Filter + Special Study Filter.*

Only use a filter to reduce a list. A filter should be cleared if you want all items regardless of that filter's values. For example, do not select anything in the Worklist Task Type filter if you want all data regardless of whether it is in the worklist.

The screenshot shows a web interface for defining search criteria. At the top, there are buttons for "Run (Ctrl+Enter)", "Reload from Last Save", and "New Patient Set Search". Below these is a table with columns for Conjunction, Field, Operator, and Value. The criteria are as follows:

Conjunction	Field	Operator	Value	clear
	Deleted [PAT]	Is	no	remove
And	Date of Last Contact Year [PAT] (#1750)	Is Less	2022	remove
And	Vital Status [PAT] (#1760)	Is	1	remove
And	Deleted [CTC]	Is	no	remove
And	Age at Diagnosis [CTC] (#230)	Is Less or Equal	19	remove
And	SEER Reportable Status [CTC]	Is	1	remove
And		Is		

The statement above will return Patient Sets that meet these criteria:

- Not deleted
- Date of Last Contact prior to 2022
- Vital Status = Alive
- And have at least one CTC that meets these criteria:
  - CTC is not deleted
  - Age at DX <= 19
  - SEER Reportable Status = 1 (it is SEER reportable)

## Using Auto-complete to Search for a Field

An auto-complete search is used to find fields for the Field Criteria, Display Fields, and Order By features. There are several fields that are stored in different types of data. For example, last name is a field in Patient Set Informants, Patient Set Aliases, and the Patient Set itself. Be careful to verify the data entity when you select a field.

The screenshot shows the 'Criteria' search interface. At the top, there is a blue header with '<< Criteria' and a '>>' button. Below the header are two buttons: 'Run (Ctrl+Enter)' and 'Reload from Last Save'. The main area contains a table with columns for 'Conjunction', 'Field', and 'Value'. The first row has a dropdown menu with 'And' selected, a field input containing 'name las', and a dropdown menu with 'Is' selected. A dropdown menu is open below the field input, showing three options: 'Name--Last [PAT ALIAS] (#2280)', 'Name--Last [PAT INFORMANT] (#2394)', and 'Name--Last [PAT] (#2230)'. There are also 'clear' and 'remove' buttons on the right side of the interface.

## Enter a question mark to view the lookup values for a field.

The screenshot shows the 'Criteria' search interface. At the top, there is a blue header with '<< Criteria' and a '>>' button. Below the header are three buttons: 'Run (Ctrl+Enter)', 'Reload from Last Save', and 'New Patient Set Search'. The main area contains a table with columns for 'Conjunction', 'Field', 'Operator', and 'Value'. The first row has a dropdown menu with 'And' selected, a field input containing 'SEER Reportable Status [CTC]', a dropdown menu with 'Is' selected, and a field input containing '?'. A dropdown menu is open below the field input, showing two options: '0 : Not SEER Reportable' and '1 : SEER Reportable'. There are also 'clear' and 'remove' buttons on the right side of the interface.

## Worklist Task Type

Worklist Task Type filter allows you to search for Patient Sets or records that are the focus of a worklist task. You would not use the Data Search just to find Patient Sets or records in the worklist. You should use the worklist for that. Use this feature to find data that meet other criteria and are in the worklist; or data that meet the criteria but are not in the worklist.

This filter only considers the Patient Set or record that is the focus of the task. Patient Sets or records listed as possible matches in Match tasks are not considered by this filter. Records can be the focus of these tasks:

- Match or Match Review
- Pathology Screening or Screening
- QC Record
- Resolve Record Errors
- Review Path Matching
- Review Pull List
- Review Special Studies

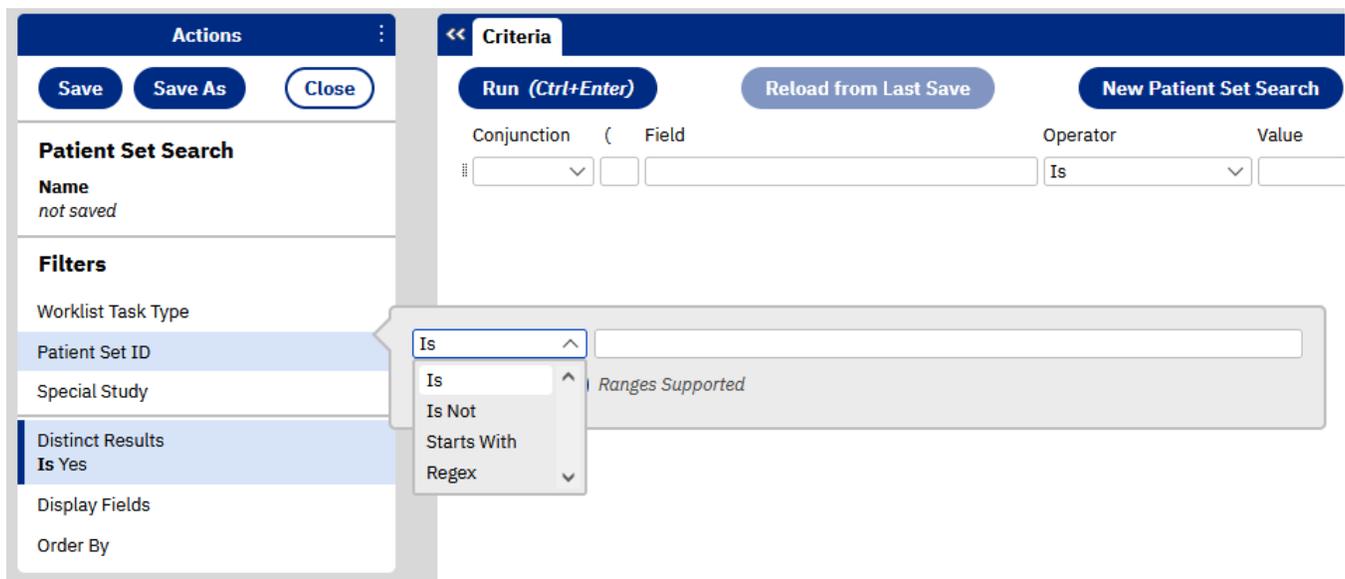
You can search for "Is Missing" to find records that are not in the workflow. You can search for "Is Not Missing" to search for records that are in any type of task.

In a Patient Set search, the Worklist Task Type filter allows you search for Patient Sets that meet other search criteria and are in the worklist; or you can use this filter to exclude Patient Sets that are in the worklist. Patient Sets can be the focus of these tasks:

- Consolidate
- QC Patient Set
- Resolve Duplicate Patient Set
- Resolve Patient Set Errors
- Review Data

One common use of this filter in a Patient Set Search would be to search for Patient Sets that are failing an edit but are not in Resolve Patient Set Errors tasks. (To search for Records or Patient Sets with a particular edit, set the Criteria to Edit Rule ID is <Edit ID>.)

### ID Filter



Use the ID Filter to search for Patient Sets or records by ID.

- You can search for a single ID or paste a long list of IDs into the filter. Each ID in a list must be separated by a blank or a comma.
- The PAT- and REC- prefixes are optional.
- Use "Is Not" to exclude certain Patient Sets or records from your search.
- The results of the ID filter will be "and-ed" with the Criteria statement and the other filters.

### Special Study Filter

Use the Special Study Filter to search for Patient Sets or records by study.

- You can search for records or Patient Sets that are in a certain special study.
- Or use "is not" to search for data that are not in a specific study.
- Or search for data not in any study (use the Missing option).

### Import ID Filter

Use the Import filter to search for records loaded in a specific import. This filter is only available when Search Type is set to Records. All imports are listed when you first click Import. Use the

search box at the bottom to find the Import ID. The search text will be matched against Import ID, File Type, Username, and Facility ID.

Complete the filter statement of "Import <is or is not> Import ID". To do this, you may enter an Import ID into the text box at the top; or check the box next to the import in the search list. You may enter a set of IDs or even a range of IDs. The IMP- prefix is optional.

## Display Fields

The Display Fields option defines the columns shown in the results table. The default list of Display Fields is defined in the registry configuration; some registries only include patient level demographic fields in the default set while other registries include tumor-level fields. To change your registry's default display fields, contact IMS.

## Order By

Add fields to the Order By list to define the sort order of results. The default sort order is Patient Set or Record ID in ascending order. You may specify multiple sort fields and you may specify ascending or descending for each. Your results can only be sorted by fields that are listed as Display Fields. Use the auto-complete search box to find a field. Click *ascending* to toggle to *descending* and vice versa.

Once a list is created, you may grab a field to move it up or down in the list.

The screenshot shows a user interface for configuring search results. On the left, there are two main sections: 'Distinct Results' with a sub-section 'Is No' and 'Display Fields' which lists 'documentId', 'reportability', 'dateOfDiagnosisYear', 'primarySite', 'laterality', and '9 more...'. Below this is the 'Order By' section, which is currently active. It displays a list of fields to be sorted by: 'Date of Diagnosis Year [REC] (#390)' with 'ascending' and 'remove' options, and 'Primary Site [REC] (#400)' with 'ascending' and 'remove' options. A search box below the list contains the text 'docum|' and a dropdown menu shows 'Document ID [REC]' and 'Rpt Cat - Report documents metastatic cancer site [REC]'.

## Convert to SQL Search

To view the SQL code used for the search, click **Actions > Convert to SQL Search**. The SQL is shown in the SQL tab. The SQL auto created by SEER\*DMS is difficult to read, but it is one way to see how the tables are linked and to see the names of database fields used in the search. It is possible, but not recommended, to use this as a way to build a query. There are SQL samples that would work better as a starting point for most queries.

## SQL Data Searches

Writing SQL Data Searches is an advanced feature that should be restricted to users who understand relational databases and have experience writing SQL queries (the *data\_search\_write\_sql* permission is required).

However, many users should be allowed to run SQL data searches. In general, a data search can be written by one person but executed by another. The *data\_search\_run\_sql* permission allows users to execute SQL Data Searches made available to their account. They may have written the data search, or it may have been written by someone else and made available to them.

<input type="checkbox"/> Name	Type	Available To	Created By	Modified By
<input type="checkbox"/> search-1.one.user.other.than.author	SQL	adamss	coyle	coyle
<input type="checkbox"/> search-2.available.to.users.in.a.role	SQL	<b>Editing Staff</b>	coyle	coyle
<input type="checkbox"/> search-3.available.to.author	SQL	coyle	coyle	coyle
<input type="checkbox"/> search-4.available.to.all.users	SQL	<b>All Users</b>	coyle	coyle

The screen shot above shows four possibilities.

- Search 1 was written by one user for another user.
- Search 2 was made available to users in a role.
- Search 3 was written and is used by the same person.
- Search 4 is available to all users.

Making a data search available to some, but not all, users can reduce clutter in a user's list of data searches. It is not intended to and should not be used to limit a user's access to data.

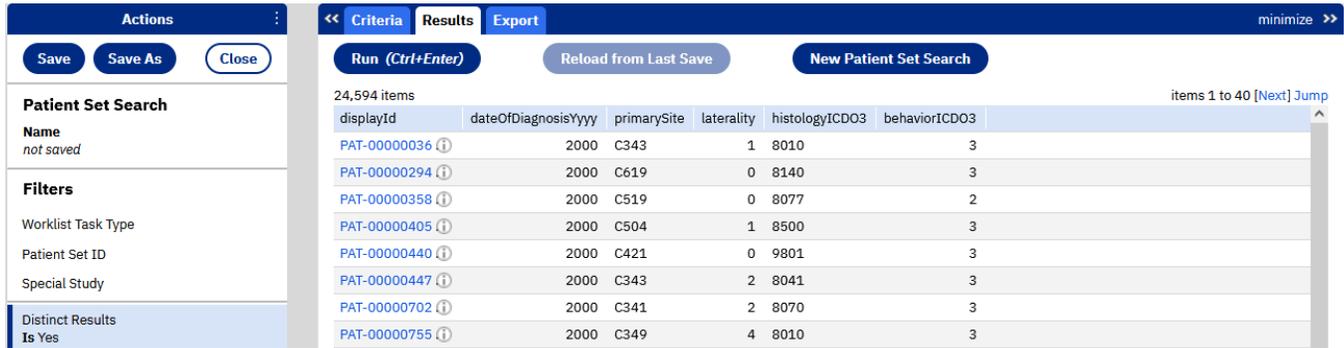
## Running or Changing Data Searches

The **Run**, **Reload from Last Save**, and **New Data Search** buttons are shown above the results.

- Click Run to execute the Data Search (or press Ctrl+Enter on your keyboard)
- Use Reload from Last Save to undo your changes.
- Click the New search button to start a fresh data search of that type. The button will say New Patient Set Search, New Record Search, or New SQL Search

## Reviewing Results in the Data Search

The data returned by your search will be displayed in the Results section of the screen.



displayId	dateOfDiagnosisYyyy	primarySite	laterality	histologyICDO3	behaviorICDO3
PAT-00000036	2000	C343	1	8010	3
PAT-00000294	2000	C619	0	8140	3
PAT-00000358	2000	C519	0	8077	2
PAT-00000405	2000	C504	1	8500	3
PAT-00000440	2000	C421	0	9801	3
PAT-00000447	2000	C343	2	8041	3
PAT-00000702	2000	C341	2	8070	3
PAT-00000755	2000	C349	4	8010	3

- The total number of results is shown in the Results tab in the lower half of the screen. Click **maximize** to enlarge the viewing area.
- 40 results are displayed per page. There are page controls on the right side of the screen.
- You may open a Record or Patient Set by clicking its ID.
- Double-click any data cell to view the text in a pop-up window. This is useful for large text fields and JSON fields, although any field can be enlarged. You may also copy the text to your clipboard.
- Click the double arrow next to the word Criteria to hide the filter panel and expand the width of the results table. Click it again to re-open the filter panel.

## Exporting Data Search Results

You can export the results to text files in comma separated values (CSV) format. The file can be saved to disk or opened in Excel. Note: Excel automatically drops leading zeroes and can create issues when processing exported data with other software. Avoid opening and saving via Excel if the original formatting is needed.

- **Export** – Use the **Export** option to create a text file containing all rows and columns in the results. This option is only available to users with the *data\_search\_export* permission. The export file will be a Comma Separated Values (CSV) text file, a common file format that can be opened in Excel and other spreadsheet applications.
- **Export to Report Output Task** – The report will be generated as a Report Output Task. Click the TSK-ID provided in the Export Tab to view the output. The format will be a CSV text file and will be identical to the export file created using the **Export** option.

# Saving and Closing Data Searches

The **Save**, **Save As**, and **Close** buttons are in the left panel.

- Use the Save button to overwrite the current data search with your changes and keep the same name. (Alt + S also works)
- Use Save As to create a new version of the data search.
- Use Close to return to the Data Search Manager

**Save Search As**

**Search Information**

Overwrite \*

Name \*

Available To \*      User \*  
     

Description

---

**Add as Manager Page Filter** ⓘ

Display ID Column

Display ID Type

Manager

SaveCancel

## Search Information

The Overwrite drop down can be set to create a new data search or the name of a data search to be updated.

Select the Available To option. If limiting to a user or role then you will be prompted for that information.

## Add as Manager Page Filter

Data Searches may be used as a filter in the Worklist, AFL Manager, and other managers.

Your Data Search must return an ID that is shown on that Manager Page. Click the information icon (next to the word Filter) for specific instructions.

This is an advanced option. Use with care. It can be a convenient way to apply the data search to results returned in manager pages, but it only works well if it is a limited or efficient query that runs quickly.

## Creating Worklist Tasks via the Data Search

You may use the data search to create worklist tasks for Patient Sets or records. You may also use the data search to send records to a step in the workflow that is earlier than their current workflow task. These are advanced options that should be used with care. Registry staff should request recommendations and support from the IMS technical support team.

### Create QC Tasks

A QC Task can be created for each record or Patient Set in the results. You will be prompted for the following information:

- Assign To – You may assign the tasks to another user, your own account, or leave them unassigned.
- Flag – You have the option of setting a flag for the tasks as you create them.
- Subject – This is the name of the QC project. It will appear in the worklist information column.
- Instructions – Enter instructions for the coders who will complete the tasks.
- Parameters – This feature is only available to IMS users. Use the parameters tab to create fields that will be shown in the QC Task. If you need QC Tasks with parameters, please submit a request to the SEER\*DMS technical support Squish project.

### Create Resolve Patient Set Error Tasks

At times, there may be Patient Sets failing edits that are not in the workflow. This happens when edits are modified, and the new logic creates edit failures for Patient Sets that completed the workflow. Patient Sets are not automatically sent into the workflow every time an edit is changed because that could create unnecessary tasks if a mistake were made when modifying an edit.

Use Create Resolve Patient Set Errors Task to send Patient Sets into the automated edits task of the workflow. A Resolve Patient Set Errors task will be created for each Patient Set that is failing one or more edits.

### Create Annotation Tasks

This option can be used to create Annotation Tasks for records. The annotation project must be created in advance by the IMS support team, therefore, submit a tech support issue if you wish to use this feature.

The Annotation Task is a manual task to code project-specific data items based on text. A registrar codes data items while reading text-based reports and highlights supporting text for each data item. To date, the SEER\*DMS annotation tasks have been used to support NCI quality control and development projects.

### Send Records into Workflow

This option sends records to an earlier point in the workflow. If you are unsure of the appropriate entry point or possible ramifications of sending these records into the workflow, please contact SEER\*DMS Technical Support. It is recommended that you test the process on the SEER\*DMS test server before sending a large number of records into the production workflow.

## SQL Tools

Tools for writing SQL are available in the Actions menu. These are listed below and are described in greater detail on the Tools tabs of the SQL Data Search.

- **Create SQL List** – Converts an unformatted list into SQL syntax that can be used with an IN statement. Sample result:  
*('REC-1', 'REC-2', 'REC-3')*
- **Create SQL List – Multiple Fields** – Converts a delimited list into SQL syntax for multiple fields. Sample result:  
*((display\_id = 'PAT-00000001' AND tumor\_record\_number = '01') OR (display\_id = 'PAT-00000002' AND tumor\_record\_number = '02') OR (display\_id = 'PAT-00000003' AND tumor\_record\_number = '03'))*
- **Define Parameters** – Use this tool to prompt users for information when they run the data search. For example, you could write a generic SQL statement that can be used year after year by prompting the user to update parameters for years of diagnosis. You can define a parameter as optional or required. The parameter will be displayed in the left panel. If you would like to set a default value, then enter the value for the parameter and save the data search. Instructions for creating parameters are provided on the Tools tab.
- **View Lookups** – You may use a SEER\*DMS lookup to associate valid values with a parameter. The parameter will show a value and a label. View Lookups to see the SQL and possible values for the lookups. The first field in the SQL will be shown as the value in the parameter; the second field will be shown as the label.
- **View Procedures** – Use this tool to view a searchable list of stored SQL procedures. Click on the procedure to view its arguments and source code.
- **Explain Query** - This displays the PostgreSQL execution plan for the SQL statement. The execution plan shows how tables referenced by the statement will be scanned by plain sequential scan, index scan, etc. If multiple tables are referenced EXPLAIN shows what join algorithms will be used to bring together the required rows from each input table. The most important and useful information provided by EXPLAIN is the estimated execution cost.

## Database Documentation

The Database tab of the SQL Data Search includes a table and field search tool. To find a field in any table, enter search text in the search box. To view all fields in a table, select the table using the drop-down menu. Select a field to view detailed field documentation including the NAACCR documentation for the field.

The Help tab of the SQL Data Search includes diagrams for core SEER\*DMS tables and the relationships between those tables. Additional database documentation will be made available upon request.

## Direct Editing of Records and Patient Sets

SEER\*DMS enables users to edit records or Patient Sets accessed via the Patient Lookup or Data Search. Read/write access to data is not restricted to workflow tasks. This type of editing is called "Ad Hoc Editing." This creates the possibility that two users may open the same record or Patient Set simultaneously. It is also possible for you to open a record or Patient Set that is currently being processed in an automatic task.

If you and another user open the same Patient Set, the second person to open it will see a warning. The first one to save data will be able to successfully save their changes. If the other person attempts to save changes, SEER\*DMS will display an error message indicating that the record or Patient Set has been modified since the data were opened in the editor. SEER\*DMS will post an error message. Consider canceling out of the Patient Set and allowing the other user to complete their work; or use the Undo Changes menu option to reload the patient's data and review the other user's changes.