Chapter 24: Creating Reports and Extracting Data

SEER*DMS includes an integrated reporting and extract module to create pre-defined system reports and extracts. Ad hoc listings and extracts can be generated in the SEER*DMS Data Search, worklist, or manager pages. To create custom reports that are more complex, you may use external software such as Groovy or Perl scripts, Microsoft Access, Crystal Reports, SAS, SQL Workbench, or SEER*Stat.   SEER*Stat may be used to generate reports or analyze data extracted from SEER*DMS; the other tools may be configured to query a SEER*DMS database directly.

Complex or inefficient queries to the live database may affect system performance and affect interaction with the system by registry staff. SQL queries defined in the SEER*DMS Data Search should be written and tested in external software prior to execution in the production system. For all external extracts and reports, it is strongly recommended that you access the data warehouse unless the analysis absolutely requires data from the live database. See *Chapter 2: Records and Patient Sets* for descriptions of the live database and data warehouse.

In this chapter, you'll learn about

- Pre-defined System Reports and Extracts
    - Finding a Report or Extract
    - Organizing Reports
    - Running a Report or Extract
    - Viewing Report Output
    - Downloading an Extract Data File
    - Sending Report Output to another User
    - Viewing the SQL for a SEER*DMS Report
- Tools to Create Custom Listings and Extracts

Pre-defined System Reports and Extracts

Commonly used reports and extracts are made available within the SEER*DMS report module. Any report or extract can be implemented as a system report if it meets two criteria: 1) the fields used in the logic are available in all registry databases; and 2) it is a general purpose report or extract that will be used repeatedly.

The business requirements for system reports and extracts are proposed by registry staff or by the SEER*DMS development team. The requests are often reviewed by other registries via the SEER*DMS Technical Support Squish project.   All system reports and extracts are implemented and maintained by the SEER*DMS development team.

The data for most reports are obtained through SQL queries; however, data for some reports are obtained via Java routines. The queries for SQL-based reports can be viewed in the SEER*DMS System Files. The results are formatted using report shells and templates developed in the Jaspersoft iReport Designer (www.jaspersoft.com).

The *reports* system permission is required to access the Report Manager. Therefore, a user must have the *reports* permission to run any report or extract. Reports that provide information related to user productivity also require the *reports_management* permission.   A report's permission is displayed on its Report Specifications page. Extracts are only listed in the Report Manager if the user has the *extract_create* permission.

Finding a Report or Extract

System permissions: *reports*; *reports_management* and *extract_create* may also be required.

To find a SEER*DMS Report or Extract:

1. Select **View** > **Reports**.

   a. Reports and extracts are organized into categories. These include a category for reports that you used recently, pre-defined categories, and categories that you create. When you access the Report Manager, SEER*DMS will display the category that you viewed when you last visited the Report Manager. If this is your first time accessing reports, all reports and extracts will be displayed.

   b. Reports and extracts are listed by ID. Report IDs have an "RPT" prefix; extract IDs use an "EXT" prefix.

2. To view a list of reports in a category, click the category name displayed on the left.

3. To search for a report by ID, title, or description, type text in the box labeled **Type to Search** and press the Enter key.
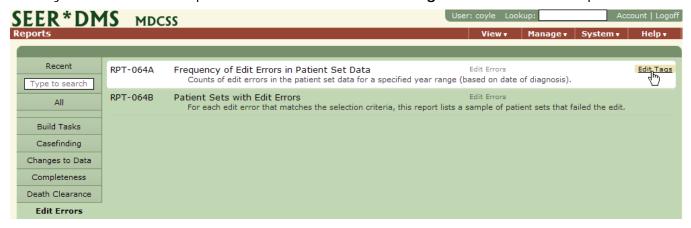
Organizing Reports

System permissions: *reports*

Report categories are defined by tags associated with the reports. A report's categories or "tags" are listed next to its title in the Report Manager. You may create new categories by adding tags. The tags that you create will be visible to all others who use the report manager.

To organize reports into your own categories:

1. Click **View** > **Reports.**

2. Use the search box or category links to find the report.

3. Hover your cursor over a report to select it. Click the **Edit Tags** link shown for the report.
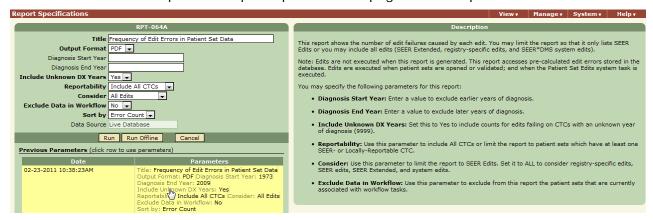

4. To create a new category, click **Specify New Tag**.

5. To add the report to a category that you created previously, click the appropriate tag in the Add Tag list. You may not add a report to a pre-defined category.

6. To remove a report from a category, click the category's tag in the Remove Tag list.

Running a Report or Extract

System permissions: *reports*; *reports_management* and *extract_create* may also be required.

*To run a pre-defined system report or extract in SEER*DMS:*

1. Select **View** > **Reports** to access the Report Manager. Enter search text or open one of the categories on the left to find the report.

2. Click the ID or title to open the Report Specifications page for the report or extract.

3. Use the Report Specifications page to define the report's output format and parameters.

    a. The report parameters are described in the **Description** section of the page. If you would like to view the SQL of the report, refer to the instructions in the *Viewing the SQL for a SEER*DMS Report* section of this chapter.

    b. Specifications used previously are available in the **Previous Parameters** section of the page. Click a row to set all specifications to values used in a previous run.

4. Edit the **Title** as necessary. This text will be listed in the Worklist's Information Column for the Report Output task. For some reports, this title will be shown as the first line in the report itself. The titles of other reports are automatically set to display parameter settings and, for those reports, the title entered here will not be included in the output.

5. If this is a report, set the **Output Format** to PDF or CSV (Comma Separated Values) based on the following considerations.

    a. **PDF:** Forms and record printouts should be output to PDF and may not display properly in CSV format.

    b. **CSV:** Many of the reports are tabular and can be exported to Comma Separated Values (CSV) format. You may open a CSV file in Excel or in a text editor. The CSV format is recommended if you wish to copy Patient Set or Record IDs from the report and paste them into the Lookup or Data Search (see *Chapter 20: Searching for Records and Patients*). Tables that do not have consistent rows and columns may not export correctly to CSV.

6. Use the parameters to customize the report or extract. If a parameter's label is in bold text, it is required.

    a. A blank value for an optional parameter implies "all possible values" for that field.

    b. Some reports have the potential to request large volumes of data from the database. To prevent a report from affecting the response time for other users, SEER*DMS terminates report queries that exceed the time limit or generate a large result set. The default thresholds are defined by the *reporting.time-out.minutes* and

reporting.max.rows configuration settings. SEER*DMS enforces these maximums unless a report-specific value is specified in the report's XML file. If a report exceeds the time or row limit, the Report Output task will display an error message. You will then have the opportunity to select new parameters to reduce the amount of output.

7. Run the report or extract by clicking the appropriate button:

   a. **Run** – queries the database and creates the report as you wait. (This option is not available for extracts.) If this process is able to complete within the report's time limit, the Report Output task will open automatically. If the time limit is exceeded, the report will automatically be run offline and the Report Output task will be available in the worklist when the task completes.

   b. **Run Offline** – creates an automated, Run Report workflow task. You will be returned to the Report Manager.

8. To open the report, follow the steps in the *Viewing Report Output* section of this chapter. To download the extract data file, follow the steps in the *Downloading an Extract Data File* section of this chapter.

In order to protect the system from a slowdown caused by a large query, SEER*DMS limits the number of report and extract tasks that are processed simultaneously. Additional report and extract tasks are queued and are executed when others complete.

Viewing Report Output

Requires system permission: *reports* (certain reports also require *reports_management*)

An automated Run Report worklist task is created when a report is executed. When the report is successfully generated, a manual Report Output task is created and assigned to the user who submitted the report. If the report was executed using the Run command, this Report Output task opens automatically (proceed to step 2 in the instructions below). If a Run Report task fails, a System Failure task is created. SEER*DMS users with the *system_administration* permission can access and review System Failure tasks.

SEER*DMS restricts the execution time and number of result rows generated by each report. The default thresholds are defined by the *reporting.time-out.minutes* and *reporting.max.rows* configuration settings. These defaults can be over-ridden by a setting in the report's XML. If either limit is exceeded, an error message will be displayed. You may modify the specifications to reduce the size of the report and re-execute.

To view the output of a report that you submitted:

1. To view the output from a report that ran offline:

   a. Filter the worklist to view Report Output tasks assigned to your account.

   b. If you are searching for a recently generated report, click the **Task** column header in the worklist. This will reverse the sort order and bring the most recently run report to the top of the list.

   c. To search for a report by title, enter a full or partial title into the **Information** filter and click **Apply**.

   d. If the task is not displayed in the list, the report may have failed or may still be running (see the instructions for *Searching the Worklist for Report and Extract Tasks*).

   e. Click on the Task ID to open the Report Output task.

2. Click **View**. You will be prompted to Open or Save the report. Click **Open**. If the output format is PDF, open the report with Adobe Acrobat. If the output format is CSV, open the report with Excel.

   a. Use the Adobe or Excel controls to print or save this report.

   b. Close the Adobe or Excel window.

3. If you need to correct parameter values or titles, submit a new report:

   a. Click **Edit Specifications**.

   b. You may make modifications to the **Title, Output Format,** and **Parameter** fields.

   c. Click **Run** or **Run Offline**.   A new task will be submitted. The existing task will remain in the worklist until it is removed. You will be returned to the worklist.

4. If you will not need to access this Report Output task again, you should remove the task from the worklist to prevent unnecessary clutter. Click **Remove** to permanently delete this task and the PDF or CSV file from the worklist. If you may need to access this Report Output task again, click **Close** to save and return to the worklist.

Report Output tasks are auto-removed based on configuration parameters (*report.removal.time* and *extract.removal.time*). The default settings are 30 days for reports and 90 days for extracts.

Downloading an Extract Data File

Requires system permission: *reports* and *extract_create*

An automated Run Extract Report task is created when a SEER*DMS extract is executed. When the task completes, an email is sent to the user who initiated the task and a Report Output task is created and assigned to the user. If an extract task fails, a System Failure task is created.

To download an extract data file:

1. Open the extract's Report Output task in the worklist.

2. Review the following items that are provided in the Report Output task:

   a. **Num Records** – the number of records included in the extract data file. If this value is zero, review the parameters that were specified to determine if an error was made.

   b. **Extract Files** – a list of one or more files related to this extract. The number of files varies by extract. You have the option of downloading one of the files or downloading a Zip file containing all of the files.

   c. **Report Specifications** – verify that the correct parameter settings were used to create the extract.

3. If you need to correct parameter values, submit a new extract:

   a. Click **Edit Specifications**.

   b. You may make modifications to the **Title** and **Parameter** fields.

   c. Click **Run Offline**.   A new task will be submitted. The existing task will remain in the worklist until it is removed. You will be returned to the worklist.

4. To download the extract data file(s):

   a. Tips for managing downloads in Firefox:

      i. Typically, files will be downloaded to a folder that is defined as the "Downloads" folder in Firefox. In the Firefox menus, go to Tools > Options and select the Downloads tab. Verify that the downloaded data will be stored in an appropriate

location. Change this location, if necessary. Please refer to the Firefox user's manual for more specific instructions.

    ii. At any time, you may access your download folder by selecting Tools > Downloads from the Firefox menus.

  b. Click **Download** or **Download All** on the Report Output Task page. Depending on browser settings, the file may be saved to your download folder automatically or you may be prompted to Open or Save the file. If prompted, click **Save to Disk**. If you are not prompted to select a location, the file will be saved in the folder that is defined as the "Downloads" folder in Firefox.

5. If you will not need to access this Report Output task again, click **Remove** to delete the task from the worklist and delete the data file from the server. If you may need to access this Report Output task again, click **Close** to save and return to the worklist.

Sending Report Output to another User

Requires system permission: *worklist_task_reassignment*

Typically, you should distribute report and extract data by saving the files to disk or providing printed copies of reports. However, SEER*DMS does allow you to reroute Report Output tasks to other users via the same mechanism used to reroute any worklist task.   Users must have the appropriate permissions to access the output tasks.

To reroute an open Report Output task, select Reassign from the **TSK** menu in the task. This is the task menu and is displayed as the Task ID on the menu bar next to View. Set the **Action** to *Assign to User Account*, and select a user name from the list. The list will only include users who have the system permissions required to open the task.

You may also use the modify feature of the worklist to reroute a Report Output task or a set of Report Output tasks (instructions for using the modify feature of the worklist are provided in *Chapter 4: Using the Worklist*).

Viewing the SQL for a SEER*DMS Report

Requires system permission: *system_administration*

XML files containing source code for system reports are included in the System Files segment of the System Administration page. If a report is implemented in SQL, the query will be included in its System File.

*To view a SEER*DMS Report's SQL:*

1. Click **System** > **Administration**. The current server log will be shown by default.

2. Select **System Files** from the **Module** drop-down list.

3. Select the report's XML file from the Reports section of the **Files** drop-down list. Report filenames begin with the report ID (e.g., *rpt064A.patient.edit.error.summary.cfg.xml*).

4. Search for the "query" XML tag. If the query type is SQL, the SQL will be displayed within the XML file. If the query type is Java, the source code will not be displayed.

5. SQL for system reports include non-SQL code to implement report parameters. Refer to the Sample SQL section of the SEER*DMS Web Portal for instructions in modifying the SQL so that it can be used in external applications like SQL Workbench.

Tools to Create Custom Listings and Extracts

In SEER*DMS, manager pages are screens that allow you to view, sort, and search within a specific set of data. For example, the Worklist is a manager for tasks. Other managers in SEER*DMS include the Contact List, Facility List, AFL Manager, and Death Clearance Manager.

Any data shown in the Results section of a manager page or the Data Search can be exported to a CSV file.   The CSV file can be downloaded and used as input to an external program; or it can be opened in Microsoft Excel or other spreadsheet software.   Use the controls in Excel to format and print the listing.

System Permissions to Export Data: Any user who can access a manager page can export the results from that page; the *data_search_export* permission is required to export data from the Data Search.

On a manager page, use the filters to define the data set. A tutorial for using the filters is available on the SEER*DMS website at seer.cancer.gov/seerdms/manual.   Instructions are also provided in *Chapter 3: Using SEER*DMS*.

The look-and-feel of the Data Search is similar to other SEER*DMS filters. However, the Data Search includes specific tools for setting search criteria and choosing display fields. Instructions for using the Data Search are included in the *Search Tools* tutorial available on the SEER*DMS website and in *Chapter 20: Searching for Data*.