Chapter 24: Reports and Extracts

SEER*DMS includes a built-in reporting module that allows you to generate standard reports and data extracts. IMS develops and maintains system reports and extracts based on business requirements submitted by registries. Together with operational dashboards, these reporting capabilities help monitor completeness, timeliness, productivity, data quality, and other key aspects of registry operations.

You can also use the Data Search module to quickly generate ad hoc listings and export results in CSV format, giving you flexible access to registry data (see Chapter 20).

For more complex reporting and analysis, external tools such as SAS, SQL Workbench, Tableau, and SEER*Stat can be used. These tools can analyze data extracted from SEER*DMS or query the database directly to support custom reports, advanced analytics, and specialized data extracts.

In this chapter, you'll learn about

- Pre-defined System Reports and Extracts
  - Reports & Extracts Page
  - Running a System Report or Extract
  - Viewing Report Output
  - Downloading Extracts
  - Sending Report Output to another User
  - Viewing the SQL for a SEER*DMS Report
- Creating a Custom Extract
- SEER*DMS Dashboard Reports & Extracts

Pre-defined System Reports and Extracts

Commonly used reports and extracts are available within the SEER*DMS Reports & Extracts module. Reports or extracts intended for repeated, general use should be implemented as system reports or system extracts. To request a new report or extract, submit a request through the SEER*DMS Technical Support system.

Requests for new reports or extracts may be submitted by registry, IMS, or NCI staff. Proposed requirements are typically reviewed through the SEER*DMS Technical Support Squish project, where registries can provide feedback. When appropriate, parameters are added so a single implementation can support multiple registries.

Most reports retrieve data using SQL queries, although newer and updated reports are increasingly implemented in Java. If you have questions about the logic of a system report or extract, please submit a technical support issue. Report results are formatted using report shells and templates developed in the Jaspersoft iReport Designer.

The system permissions related to reports and extracts are described below. The system permission required for each report is listed on the Reports & Extracts page and on the Report Specifications page.

- *reports* - Run and view reports that do not require the *reports_management* permission.
- *reports_management* - Run and view all reports, including management reports. These include reports showing staff productivity and user access information.

- *reports_support* - Run reports containing user information (but not patient data). Intended for registry IT staff who support users but do not have permission to view protected patient information.
- *extract_create* - Create a data file by executing one of the standard or custom extracts from the **Reports & Extracts** page.
- *manage_custom_extracts* - Create, edit, or maintain custom fixed-column extracts using the Extract Editor. These are available on the **Custom Extracts** tab of the **Reports & Extracts** page.

Reports & Extracts Page

The Reports & Extracts page can be accessed by selecting View > Reports & Extracts. Access to this page is limited to users who have one of the system permissions described in the previous section.

The Reports & Extracts page allows users to locate, filter, and run pre-defined reports and data extracts within SEER*DMS. Three types are available: system reports, system extracts, and custom extracts.

- IMS maintains all system reports and extracts as part of SEER*DMS application updates.
- Custom extracts are implemented using Groovy code maintained within the SEER*DMS interface. IMS or registry users can modify custom extracts. Custom extracts are limited to fixed-column data files. See the Custom Extracts section of this chapter for more information.

Reports and extracts are identified by unique identifiers:

- Reports use the prefix RPT-
- Extracts use the prefix EXT-
- Custom extracts are identified by Title.

Columns Shown on the Reports & Extract Page

The following columns are shown on the Reports & Extract Page. Click a column header to sort the results by that field. Filters are available for all displayed columns.

- ID — The unique report or extract identifier beginning with RPT- or EXT-.
- A (Access) The access level required to view and run the report or extract. Hover over the icon to view the permission associated with each access level.
- Title — The name of the report or extract.
- Category — The category to which the report or extract belongs.
- Description — A summary of the report output or extract contents. You may filter results using all or part of the descriptive text.
- Last Run — The date you most recently ran the report or extract.

Report Categories

To make it easier to find a report or extract, they are organized into the categories listed below.

- Changes to Data
- Completeness

- Death Clearance
- Duplicates
- Edits
- Extracts
- Follow-up
- Imports
- Productivity
- Quality Control
- SEER Annual Report
- System Management
- Worklist

Running a System Report or Extract

System permissions: *reports*; *reports_management, extract_create*, or *reports_support*

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

1. Select **View** > **Reports & Extracts** to access the Reports & Extracts page. Use the filters to search for a specific report or extract.
2. Click the ID to open the Report Specifications page for the report or extract.
3. Use the left panel of the Report Specifications page to set the report's parameters.
   a. The report parameters are described in the **Description** tab in the right panel.
   b. Parameter settings used previously are shown in the searchable **History** section of the page. Click a row to set the parameters to values used previously.
4. You may enter an optional Subtitle. This will be shown in the information column of the worklist for the Report Output task. For some reports, it will also be shown in the report itself.
5. If this is a report, set the **Output Format** to PDF or CSV (Comma Separated Values) based on the following considerations.
   a. **PDF:** Use PDF for formatted reports such as forms or record printouts. PDF format is not available for some reports that were designed to provide a data listing.
   b. **CSV:** Use CSV for tabular reports that will be opened in Excel or imported into other applications. CSV output is also recommended if you wish to copy Patient Set or Record IDs from the report and paste them into the Lookup or Data Search. Reports requiring formatted layouts are not available in CSV.
6. Tips for setting report parameters:
   a. If a parameter's label is in bold text, it is required.
   b. Leaving an optional parameter blank includes all values for that field.
   c. Parameter settings determine how much data is included in the report. Some reports can return very large data sets. To protect overall system performance, SEER*DMS enforces execution time limits. Reports that exceed the time limit are terminated, and

the Report Output task displays an error message. If necessary, submit a SEER*DMS Technical Support request to increase the timeout for a specific report.

7. Run the report or extract by clicking the appropriate button:
   a. **Run** – generates the report immediately while you wait. (This option is not available for extracts.) If the report completes within 30 seconds, the Report Output task will open automatically. Otherwise, the report will automatically be run offline, and the Report Output task will be available in the worklist processing is complete.
   b. **Run Offline** – creates an automated, Run Report task. You will be returned to the Reports & Extracts page.
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.

SEER*DMS limits the number of reports and extracts that can run simultaneously. Additional requests are queued and processed as resources become available.

Viewing Report Output

Requires system permission: *reports* or *reports_support* (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 you selected Run, this Report Output task opens automatically when the report finishes. If a Run Report task fails, a System Failure task is created. Users with the appropriate system permissions can review System Failure tasks.

To view the output of a report:

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, sort the worklist by Age with the most recent report at the top of the list (click the **Age** column header to sort; click again to reverse the sort order).
   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** to open or save the report.
3. If you need to change the parameters, report title, or output format, click **Edit Specifications**, make your changes, and rerun the report.
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 automatically removed according to the *report.removal.time* and *extract.removal.time* configuration settings. The default settings are 30 days for reports and 90 days for extracts.

Downloading Extracts

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:

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. **Number of Entities** – 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** – one or more files generated by the extract. You can download individual files or download all files as a ZIP archive.
    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. Modify the **Title** or **Parameter** fields as needed.
    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. Click **Download All** to download all files or click the name of an individual file to download a specific file.
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_action_reassign_mine*

Report Output tasks can be sent to other users via the standard worklist reassignment process. Users must have the appropriate permissions to access the output tasks.

To reroute an open Report Output task, select Reassign from the **Task** menu in the task. Set the **Action** to *Assign to User Account(s)*, and select a user name from the list. The list will only include users who have the system permissions required to open the task. Use the Options dropdown menu to limit the user list to a selected role.

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 is 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 report is implemented in SQL, the query appears within the <query> XML element. Java-based reports do not include source code in the XML.
5. SQL for system reports includes parameter placeholders that are interpreted by SEER*DMS. Replace the placeholders with literal values before running the query in the Data Search or another SQL tool.

Creating a Custom Extract

Requires System Permission: *manage_custom_extracts*

The Custom Extract module was implemented when the NAACCR Vol II layout was a fixed-column format, therefore, it was only designed for fixed-column files. To create an extract in other formats:

- Use the SQL Data Search to create a CSV extract if your logic can be written in SQL.
- Or submit a request for a system extract via the SEER*DMS Technical Support system.

Registry IT or IMS staff can define a fixed-column extract in the Custom Extract module.

Variables stored in the Groovy context (for example, context.myVariable) are available to subsequent steps during the extract process.

To create a custom extract:

1. Select "Add Custom Extract Definition" from the menu in the Reports & Extracts page.
2. Enter a **Title** for the custom extract and enter text for the **Description**. The Description is displayed in the right panel when users run the custom extract.
3. Define the **Parameters** for the custom extract. Use this tool to prompt users for information when they run the extract. For example, you could write a generic extract that can be used year after year by prompting the user to update parameters for year of diagnosis. You can define a parameter as optional or required. Instructions for creating parameters are provided on the Help tab on the right side of the page.
4. Define a **Cohort**: The Cohort tab contains the SQL that will return the data for the extract. You should write and test the SQL in the Data Search first.
5. The **Pre-Extract** tab can be used to define a Groovy Script that runs once at the beginning of the extract process. It can be used to initialize variables that may be needed later. The Pre-Extract script cannot access the cohort.

6. The **Pre-Line tab** is used to define a Groovy Script that runs once before each line of the extract is processed. It can change or add to the current row of the cohort. To access or change a value in the row, use "row.my_column". New values can be added to the row in the same way (i.e. "row.my_new_column = 'new column value').
7. The **Post-Extract** tab is used to define a Groovy Script that runs once at the end of the extract process. It can be used to clean up variables that may have been created. The Post-Extract script cannot access the cohort.
8. Up to five files can be created at once in a custom extract. Use the **Files** tabs to define them, starting with the "Unnamed File" tab. Refer to the Files section of the Help text in the Custom Extract module for instructions.

SEER*DMS Dashboard Reports & Extracts

SEER*DMS dashboards provide registry managers with real-time information to monitor data quality, workflow, and workload distribution. Dashboards display information in tables, graphs, and charts that help users quickly identify trends and areas requiring attention.

Depending on the dashboard, users can:

- Download a PDF version of the report.
- Export the underlying data for additional analysis.
- Open the results in the SEER*DMS Data Search for further investigation.

Dashboard data is refreshed nightly. Users can update a dashboard at any time by clicking **Refresh**.

New dashboards are added periodically to support registry management activities. Examples include:

- **SEER Data Quality Profile (DQP):** Displays submission data quality measures based on current data. The DQP dashboard provides registry managers with real-time access to the status of their submission data.
- **Edits:** Displays the number of cases failing each edit and helps registry managers verify that all required edits are active.
- **Consolidation Tasks:** Summarizes manual consolidation tasks by data item and diagnosis year and provides a snapshot of the registry's manual consolidation workload.
- **Data by Facility:** Monitors data submissions and completeness by reporting facility.
- **Possible New Cases:** Identifies source records in the workflow that may result in new incident cases.