Chapter 15: Reference Data

"Reference" is the name of an optional tab that may be displayed in the right panel of the SEER*DMS Patient Set Editor. This tab can be used to display the results of searching for the patient's data in external data; or it can be used to display a summary of data that are in SEER*DMS database tables.

The Reference tab could be used to display data from an external source to verify the patient's address, birth date, or other fields. For example, entries in motor vehicle data that are possible matches to the patient could be displayed in the Reference Data tab. The matching criteria are defined by registry staff.

The Reference tab can also be used to show a summary of data items that are stored in SEER*DMS data structures. For example, all path report numbers on HL7 records linked to the patient set could be displayed as a quick reference.

In this chapter, you'll learn about

- Using the Reference Tab in the Patient Set Editor
- Defining Reference Data Sources

Using the Reference Tab in the Patient Set Editor

The Reference Tab shows results of queries for all active Reference Data Sources.
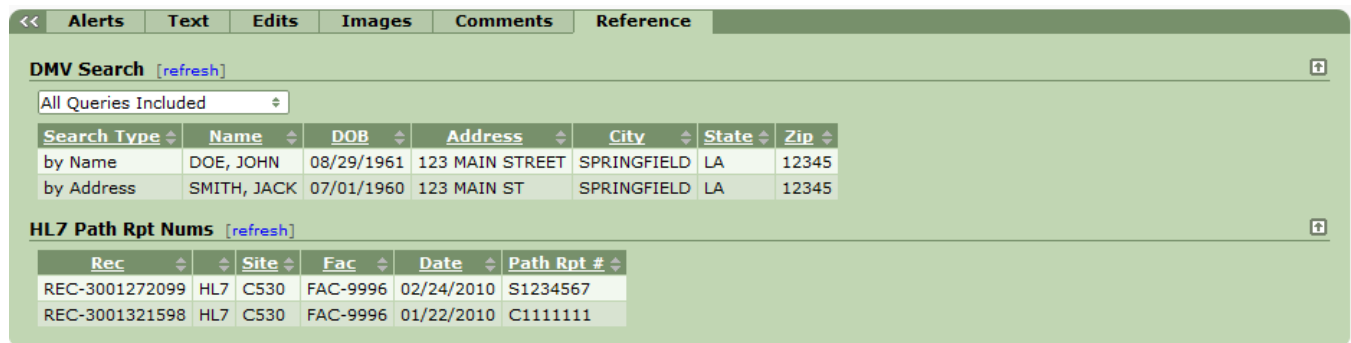
- Use the up and down arrow ⊞ to expand or collapse a Data Source section.
- If a table is too wide for the right panel, you may click the left arrow ⟨⟨ next to the Alerts tab to use the full width of the editor.
- You may click Refresh to force a query to be re-executed. You will rarely need to refresh results because the tables will automatically refresh when the Patient Set is validated or saved. Refresh is only needed if the underlying Reference data or definition are changed.
- Each Reference Data table can be sorted by clicking the column headers. Clicking the header will toggle the sort between an ascending and descending order.
- If there are multiple queries defined for a Reference Data Source, you may show or hide results from specific queries. In the example shown above, the DMV Search has multiple queries. One shows the results of a search of DMV data by name and the other shows the results of a search by street address. Results of all queries are shown, by default. Use the drop-down menu to turn specific queries on or off (shown in the figure below). Check the box next to the queries that you would like to execute. If none of the boxes are check, then all queries will be executed and displayed. The table will automatically refresh after you make your selection.
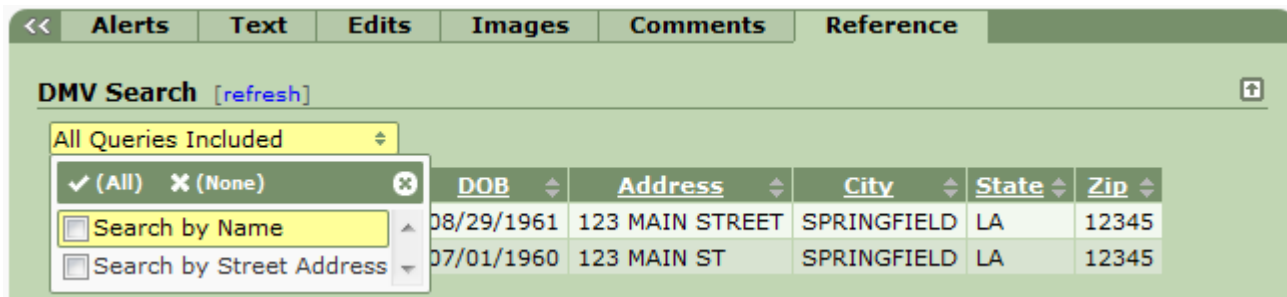
Defining Reference Data Sources

Reference Data Sources are defined by two components: the data storage and the queries that run against that data. The data may be stored in SEER*DMS database tables, the SEER*DMS warehouse, or in external data sources. External data sources include the registry's database schema (user schema).   The SEER*DMS configuration must be modified in order to support external data sources. To use data stored in external sources, please submit a request to the SEER*DMS Technical Support Squish project.

To create a new Reference Data Source:

1. Select **System > Reference Data Manager**.

2. Select **Actions > Add Data Source**

3. A Data Source template will be created.

4. Enter a **Name**. This name will be listed on the Reference tab in the Patient Set editor.

5. Verify that the **Active** box is checked.

6. The available databases are listed above the Definition text box. The production database ("seerdms") and "warehouse" are available in all registries. External data sources are available in some registries.

7. Enter a **Description** of the new data source.

8. Carefully review the XML documentation provided below and the comments in the template XML file. The XML comments provide more specific syntax. Change the queries, parameters, and result columns for your new data source. Typically, a Data Source will only have one query and you can delete the other sample. All components of the Reference Data Source XML are described below.

This is the syntax for the Data Source XML:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>


<!ELEMENT data-source (query,parameter?,result-column)>
    <!ELEMENT query (display-name, data-source)>
    <!ELEMENT parameter (name, property)>
    <!ELEMENT result-column (name, display-name)>
</data-source>
```

- **query** – At least one query is required. Multiple queries may be defined. The query element has two required attributes: display-name and data-source. The display-name is displayed in the Patient Set Editor if multiple queries are defined.   The data-source must be one of the defined data sources. The available sources are listed above the Definition text box ("seerdms", "warehouse", and registry-defined sources).

- **parameter** – This element is optional. Parameters are fields within the patient set that may be used as variables in queries. The name is the name that you are giving to this parameter. You will use ":varName" in the query source code. The property is the Java path for the field. The Java Path for each field is available in the SEER*DMS Field Mappings (click the information icon for a field on the Field Mappings help page or click the field's label in the Patient Set Editor).
- **result-column** – This element is required. You must specify at least one result-column. The name is the name of the column returned by the SQL. The display-name is the label that will be used in the Patient Set Editor.

The example definition has comments that describe each part of the XML in detail. For a complete understanding of the XML, please read those comments.

The Validate button will check that both the XML and SQL in the Definition are valid. You will not be able to save the Data Source unless both are valid.