# Table Of Contents

**Introduction to MDClone** 3  
- About Synthetic Data 3  
- User Guide Conventions 4  
- General Features and Icons 4  
- My Queries View 8  

**Building a Query Overview** 10  
- Building a Query 11  

**Defining the Patient Cohort** 13  
- Step 1: Defining a Reference Event 13  
- Step 2: Defining Additional Inclusion Criteria 16  
- Defining Events and Properties in the Query Tool 19  
  - Defining an Event Using the Search Engine 19  
  - Defining an Event with the Value Browser 22  
  - Defining an Event Manually 29  
  - Defining the Occurrence of an Event 31  
  - Defining Event Properties 35  
- Performing Additional Actions when Creating Queries 38  
  - Defining a Query with No Reference Event 40  
  - Redefining a Reference Event 40  
  - Copying and Pasting an Event Definition 41  
- Calculating the Cohort 41  

**Defining the Output Definition** 45  
- Step 3: Adding and Defining a Time-Related Event 46  
  - Setting the Timing Definition of the Time-Related Event 49  
  - Defining a Multiple Occurrence Event 53  
  - Defining a Time-Related Event with Slider 56  
  - Defining the Output Columns 58  
  - Viewing a Data Preview 61  
- Step 4: Defining Patient Demographic Columns 65  
- Performing Additional Actions on Time-Related Events 67  

**Summary Review** 70  
- Step 5: Finalizing the Cohort and Output 71  
  - Downloading Original Data 76  
  - Downloading Synthetic Data 80  
  - Adding Calculated Fields 84
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adding a Custom Expression</td>
<td>84</td>
</tr>
<tr>
<td>Performing Data Conversion</td>
<td>88</td>
</tr>
<tr>
<td>Adding Two-Column Calculation</td>
<td>90</td>
</tr>
<tr>
<td>Adding One-Column Calculation</td>
<td>92</td>
</tr>
<tr>
<td><strong>Managing My Queries</strong></td>
<td>94</td>
</tr>
<tr>
<td>Performing Additional Actions on Queries</td>
<td>94</td>
</tr>
<tr>
<td>Viewing Notifications for Event Updates</td>
<td>95</td>
</tr>
<tr>
<td>Editing a Query</td>
<td>96</td>
</tr>
<tr>
<td>Sending a Query</td>
<td>96</td>
</tr>
<tr>
<td>Sharing a Query</td>
<td>98</td>
</tr>
<tr>
<td>Duplicating a Query</td>
<td>101</td>
</tr>
<tr>
<td>Deleting a Query</td>
<td>101</td>
</tr>
<tr>
<td><strong>Creating Quality of Care</strong></td>
<td>102</td>
</tr>
<tr>
<td><strong>Managing My Events</strong></td>
<td>105</td>
</tr>
<tr>
<td>Combined Event Overview</td>
<td>105</td>
</tr>
<tr>
<td>Creating a Combined (AND) Event</td>
<td>106</td>
</tr>
<tr>
<td>Defining a Combined (OR) Event</td>
<td>110</td>
</tr>
<tr>
<td>Creating a Private Loaded Event</td>
<td>111</td>
</tr>
<tr>
<td>Updating a Private Loaded Event</td>
<td>116</td>
</tr>
<tr>
<td>Viewing My Events List</td>
<td>119</td>
</tr>
<tr>
<td>Sending an Event</td>
<td>121</td>
</tr>
<tr>
<td><strong>My Permissions</strong></td>
<td>122</td>
</tr>
<tr>
<td><strong>Importing Files</strong></td>
<td>124</td>
</tr>
<tr>
<td>Viewing Imported Files</td>
<td>124</td>
</tr>
<tr>
<td>Importing a File</td>
<td>125</td>
</tr>
<tr>
<td>File Mapping</td>
<td>127</td>
</tr>
<tr>
<td><strong>My Output History</strong></td>
<td>130</td>
</tr>
<tr>
<td><strong>Appendix A - Glossary</strong></td>
<td>132</td>
</tr>
<tr>
<td><strong>Appendix B - User Permissions</strong></td>
<td>133</td>
</tr>
<tr>
<td><strong>Appendix C - Decryption Solution</strong></td>
<td>135</td>
</tr>
<tr>
<td>Decryption Solution Overview</td>
<td>135</td>
</tr>
<tr>
<td>Getting Decrypted Data</td>
<td>135</td>
</tr>
<tr>
<td>Date Value Retrieval</td>
<td>136</td>
</tr>
</tbody>
</table>
1. Introduction to MDClone

MDClone introduces a new healthcare data paradigm, removing the barriers to access and share data in real-time and at big data scales. MDClone’s Healthcare Data Platform combines the most modern big data technologies and unique approaches to healthcare data organization, management, analysis, and sharing. Powered by proprietary technology for longitudinal data organization and for producing synthetic data, MDClone enables unprecedented insights with zero-risk to patient privacy. MDClone platform supports the Chrome and Edge browser.

Disclaimer! The figures displayed in this document might not appear the same for all users, and depend on the settings and requirements of your organization.

Disclaimer! MDClone bears no responsibility for any actions of a client undertaken in reliance on data processed by the software. Additionally, the client is solely responsible for independently validating all outputs from MDClone software.

Note that the MDClone software is not a substitute for medical advice!

This User Guide provides a detailed description of how users can build a query in the MDClone platform and perform many additional query-related tasks. It provides information regarding the following:

- Building a Query Overview
- Building a Query
- Defining the Patient Cohort
- Defining the Output Definition
- Summary Review
- Managing My Queries
- Creating Quality of Care
- Managing My Events
- Creating a Private Loaded Event
- My Permissions
- Importing Files
- My Output History

A table describing and comparing permissions granted to different users is provided in Appendix B - User Permissions.

1.1. About Synthetic Data

The MDClone engine process can examine medical data, extract their statistical characteristics and dependencies, and extrapolate from these to generate a new data set containing synthetic medical data for fictitious patients, thus overcoming the risk of public disclosure of medical data from actual patients.
1.2. User Guide Conventions

The following conventions are used in this User Guide:

- The user guide addresses users in the masculine gender but refers to both genders.
- Cross-references are marked in blue underline.
  Example: Chapter 4: Heading
- The following icons are used:

  ![Caution](image)
  **Caution:** Indicates instructions, or cautionary notes, which, if not followed, might result in damage to the data or quality of measurements.

  ![Note](image)
  **Note:** Provides helpful information and tips.

1.3. General Features and Icons

The following functions and icons that are used repeatedly throughout the MDClone Query Tool screens are described in this section. The descriptions will not be described repeatedly in the chapters in this guide.

- **Navigation Pane**

After the user logs into the Query Tool platform, the navigation pane is located on the left side of the screen. By default, after login the menu is collapsed. The user must actively open the navigation menu. From here, the user can navigate to other pages of the platform.

The **Help and Guides** option can be expanded to view the available guides. By default, the following guides are available to the end user: User Guide and Online Training Guide.
- **Query Wizard**

  While building a query, the Query Wizard is displayed in the upper part of the screen during each step in the query building process. The Query Wizard provides easy navigation back and forth between the different steps used in creating a query.

- **Notification**

  A Notification icon is provided, where relevant, in the upper right part of the screen. Clicking this icon displays notifications regarding the status of actions performed in the dashboard (for example, success or failure messages).

- **Sort**

  A Sort functionality is often provided when the data is displayed in the form of a table. Clicking the title of each column sorts the data presented in the table (alphabetically or according to date).

- **Filter**

  A Filter icon is often provided to enable the user to filter a column in a table with data. Clicking the filter icon displays a dropdown list with options that are relevant to the displayed table. The user can select which data to present by selecting the relevant checkboxes and clicking **Apply**.

- **Action Icon**

  An Action icon is often provided in screens and tables to enable the user to perform additional actions on the displayed data (for example, **Edit** or **Delete**). Clicking the icon displays a dropdown menu of available actions.

- **Search**

  A Search function is often provided, where relevant, in the upper part of screens where a table is displayed. The Search function enables the user to find the required value or text in the table. To search, type the value in the **Search** text box. The table is filtered to display all items that contain the search term. The search term in each row is also highlighted.
Details Pane:

To see details about an event and its property in an event list, move the cursor over a row in the Event list. Some or all of the following information is displayed in the Details pane to the right.

- **Group**: The name and icon of the event’s group.
- **Event Description**: A description of the event as provided by the admin. If more information was linked to the event, the user can click the More details link to access a PDF or HTML file that provides more information.
- **Property Description**: A description of the property as provided by the admin.
- **The available data is since**: The earliest year from which this data (with at least 500 rows) in the database is available.
- **% Populated**: The percentage to which this Event property is populated (with rows that are not empty) in the database.
- **Data Type**: One of the following data types (with a Type icon) is displayed: Numeric, String or Date. The default date (a date property that is fully populated) is also marked with an icon.
- **Event Type**: If the event is a Combined Event or Population Event, its type is also displayed.

Population Event:

- **Data Sample**: A sample of the data.
● **Page Scrolling**

A page navigator is provided, where relevant, at the bottom of the screen. The navigator can be used to scroll through pages for the relevant data.

● **Calendar**

A Calendar icon and function is provided in cases where the user must select a date.
1.4. My Queries View

When the user logs in to the dashboard, the My Queries screen is displayed by default. This is the landing page.

The following figure shows the My Queries screen, where the user can manage his queries and the queries that were shared with him (under Queries Shared With Me).

![My Queries Screen](image)

To view and manage My Queries:

1. Log in to the MDClone Dashboard. The My Queries screen is displayed. The My Queries screen displays the following information:
   - **MY QUERIES**: The number of queries in the list of My Queries.
   - **Name**: Name of each query. Click the name to open the Query Tool and view and edit the query.
   - **Modified**: Date the query was last modified.
   - **Created**: Date the query was created.
   - **Shared**: The Share icon indicates if the query was shared (for details, see Sharing a Query).
   - **File**: The File icon indicates if an output file was generated. A tooltip shows the date the file was created.
   - **Key**: Shows the Approval Key for the query (if one exists).
   - **Columns and Rows**: Number of columns and rows in the latest dataset file.

2. To perform additional actions on My Queries or while creating a new query, select the query on which you want to perform additional actions. Click the Action icon to display the available actions that can be performed on the selected query. The available actions might differ during different stages in creating queries.
   - **Save**: The Save option is available while creating a query either by clicking the Save icon or from the Action menu. The Save option is only available when there are unsaved changes.
3. To expand the **Queries Shared with Me** list click the arrow on the left. The Shared Queries list is displayed.

The following information is displayed for **Queries Shared with Me**:

- **Number** of queries in the list.
- **Name**: Name of the query. Click the name to open the Query Tool and either view or edit the query (depending on the granted permissions).
- **Modified**: Date the query was last modified.
- **Shared by**: Name of the user who shared the query (for details, see **Sharing a Query**).
- **My Permissions**: Permissions granted for this query, which are either **View** or **Edit**.

4. Click the **Action** icon to display the available actions that can be performed on **Queries Shared with Me** (for details, see **Sharing a Query**).

- **Edit** or **View**
- **Duplicate to My Queries**
- **Remove from List**
2. Building a Query Overview

This guide describes how to build a query using the MDClone Query Tool. It describes the main processes involved in building a query and generating the final output. The steps that are part of the query building process are displayed in the Query Wizard at the top of each page in the process.

The following chapters provide a detailed description of the steps in the query building process:

- To begin building a query, follow the steps described in Building a Query. Continue by performing the following five steps in the query building process.
- Defining the Patient Cohort: During this process the user defines the query’s cohort. The cohort includes the specific patients in the organization who meet the events defined in the first two steps of the query. The following steps are included:
  - Step 1: Defining a Reference Event: In the first step, the user defines the Reference Event. The Reference Event represents the central life event for patients in the cohort (for example, Admission to the hospital, a diagnosis of cancer, or a medical procedure) to which all other events are related in time. Patients who have this Event are the patients that are included in the cohort. For details, see: Step 1: Defining a Reference Event.
  - Step 2: Defining Additional Inclusion Criteria: In the second step, the user can further narrow the cohort by defining additional events that are included or excluded. Additional inclusion criteria can be additional event(s) and/or demographic criteria. Only patients who meet all the defined events and criteria are included in the cohort. For details, see Step 2: Defining Additional Inclusion Criteria.
- Defining the Output Definition: During this process the user defines the query’s output.
  - Step 3: Defining Time-Related Events: In the third step, the user defines the information to include in the output. This information consists of events that are related in time to the Reference Event. This step does not change the cohort size. It defines the events and data that the user wants to include in the output file for the defined cohort. For details, see Step 3: Adding and Defining a Time-Related Event.
In Step 3, it is possible to view a preview of the final output by clicking the Data Preview tab. For details, see Viewing a Data Preview.

- **Step 4: Defining Demographics**: In the fourth step, the user defines the demographic criteria to include in the output, such as gender. For details, see Step 4: Defining Patient Demographic Columns.

- **Summary Review**: During this process, the user finalizes the cohort and downloads the output file.
  - **Step 5: Finalizing the Cohort & Output**: In the final step of the query, the user finalizes the cohort, selects and applies filters on the required output columns, and adds calculated columns, if required. The user can then download/synthesize the output file according to his permissions. For details, see Step 5: Finalizing Cohort and Output.

### 2.1. Building a Query

**To start building a query:**

1. Log into the MDClone platform. The default view that appears is the **My Queries** screen.
   
   If you are currently on a different screen, in the navigation pane click **Query Tool > My Queries**. The **My Queries** main screen appears (for a description of the My Queries screen, see **My Queries View**). This is the default view.

2. On the **My Queries** page click **+ New Query**.
The empty **Reference Event** screen appears, where you can begin to define the Reference Event (*Step 1* in building a query).

3. At the top of the screen, in the **Untitled Query Name** field, type a meaningful name for the query.

4. In the **Description** field, type additional information about the query (optional).

5. To save the new query, click the **Save** icon in the upper right hand corner.

6. Proceed to build the query as defined in the following chapters:
   - **Defining the Patient Cohort**
     - **Step 1: Defining a Reference Event**
     - **Step 2: Defining Additional Inclusion Criteria**
   - **Defining the Output Definition**
     - **Step 3: Adding and Defining a Time-Related Event**
     - **Step 4: Defining Patient Demographic Columns**
   - **Summary Review**
     - **Step 5: Finalizing the Cohort and Output**
     - **Adding Calculated Fields**
3. Defining the Patient Cohort

Defining the Patient Cohort is the first stage in the process of building a query. In this step, the user defines the Patient Cohort, which is the group of patients included in the query.

Defining the Patient Cohort consists of the following steps:

- **Step 1: Defining a Reference Event**: In order to define the Patient Cohort, the user defines the query’s Reference Event, which is the main inclusion criteria of the query’s cohort. This event represents the patients’ central life event (for example an admission to hospital, a diagnosis, or a medical procedure) to which all other events will be related in time. For example, the Patient Cohort might be a group of patients over the age of 40 who have a diagnosis of atherosclerosis. Patients who have this Event are the patients that are included in the cohort. For details, see: Step 1: Defining a Reference Event. This step is performed in the first tab of the Query Wizard:

- **Step 2: Defining Additional Inclusion Criteria**: In this step the user can further define and narrow the cohort by defining additional events that are either included or excluded. Additional inclusion criteria can be an additional event (Inclusion Criteria) and/or Demographic criteria. Only patients who meet all the defined events and criteria are included in the cohort. For details, see Step 2: Defining Additional Inclusion Criteria. This step is performed in the second tab of the Query Wizard:

3.1. Step 1: Defining a Reference Event

Defining a Reference Event: This is the first step in building a query. This chapter provides a summary of how to define a Reference Event. For a more detailed description of each of the steps and options, click the provided links in this section.

To define the Reference Event:

1. Log into the MDClone platform. By default, the My Queries screen opens. Alternatively, in the navigation pane click Query Tool > My Queries.

   Click + New Query.
2. **Define the Reference Event** by selecting one of the following options:
   - **Defining Event Using Search Engine**: This section describes in detail how to use the Search function to define a Reference Event.
   - **Defining an Event Manually**: This section describes in detail how to manually define a Reference Event.
   - **Defining a Query with No Reference Event**: This section describes how to build a query without a Reference Event, by selecting the No Reference Event option from the event’s Action menu. In this case, the system uses the patient birthdate as the reference date.

3. **Filter the Reference Event by selecting the event occurrence**: When defining the Reference Event either manually or using the Search functionality, select the event occurrence in the Get field:
   - **Get the (event occurrence)**: Select which occurrence of the event from the dropdown menu. The default option is to get the first occurrence of the Reference Event, as shown in the following figure.

   ![](image1)

   The following options are available from the Get the event dropdown list:
   - **first, second, (or n occurrence)**
   - **first, second (or Nth) most recent** (by selecting the most recent checkbox)
   - **Latest or nearest**
   - **max, min, median**

   To view a description of the options, move the cursor over an option in the list. An explanation and example is displayed.

   ![](image2)

   For a detailed description of the options, see [Defining the Occurrence of an Event](#).
4. **Set** the **Age at event** by selecting the age values in the textboxes. Notice that the **Age at event** filter by default is defined as 18 to 120 but can be changed, as required. If required, the default minimum and maximum values can be pre-set by a special configuration. Define additional filters as described in **Defining Event Properties**.

5. Click the **Event filter** option to add more filters to further refine the event’s definition. After the Reference Event is defined, the Reference Event definition is displayed in the Reference Event screen.

   The example above shows a defined Reference Event of a Diagnosis of Atherosclerosis that occurred to patients between 18 and 120 years of age.

6. The Reference Event can be further defined, by selecting the event properties from the dropdown list and defining additional values (for a detailed description on how to define Event Properties, see **Defining Event Properties**).

7. To save your changes, click the **Save** icon in the upper right hand corner of the screen.

8. To calculate the cohort, click **Calculate Cohort** at the bottom of the screen (for information about calculating the cohort, see **Calculating the Cohort**).

9. For additional available actions and options when creating queries, see **Performing Additional Actions when Creating Queries**.

10. To continue to the next step (**Step 2: Defining Additional Inclusion Criteria**) click **Next** or click the **Additional Inclusion** step in the Query Wizard.

If no Inclusion Criteria are required, continue to **Step 3 - Time-Related Events** (see **Step 3: Adding and Defining a Time-Related Event**).
3.2. Step 2: Defining Additional Inclusion Criteria

Defining Additional Inclusion Criteria: This is the second step in building a query. After defining the Reference Event, the user can (optionally) narrow the cohort by defining the following kinds of additional inclusion criteria:

- Inclusion Criteria (of additional Events and Properties)
- Demographic Criteria

The relationship between the Reference Event and the Inclusion Criteria is an AND relationship.

- Inclusion Criteria: Patients who meet both the Reference Event and the defined Inclusion Criteria (if defined) are included in the cohort (multiple inclusion criteria can be defined).

  Note that Inclusion Criteria can be used to narrow the cohort by defining patients who either have or do not have an occurrence of an event (for example, patients who have had a heart attack but have NOT had bypass surgery). Notice also that the Inclusion Criteria are not related in time to the Reference Event.

- Demographic Criteria: The user can also define static demographic filters such as gender, age, deceased date, ID, and place of birth. Only patients who meet these criteria are included in the cohort.

This chapter provides a short description about how to define additional Inclusion Criteria. For a fully detailed description of each of the options provided in this step, click the provided links.

To define Inclusion Criteria:

1. Log into the MDClone platform. After defining one or more Reference Events in your query, click Next (or select the Additional Inclusion tab) to (optionally) define Inclusion Criteria events.
2. Define one or more additional events by Defining an Event Using the Search Engine or Defining an Event Manually. To select values using the Value Browser, see Defining an Event with the Value Browser.

The following example shows an event added as Inclusion Criteria.

After defining the Inclusion Criteria, the cohort is defined with patients who meet both the Reference Event and the defined Inclusion Criteria (if defined). The first Inclusion Criteria section is labeled 1. Additional Inclusion Criteria are labeled 2, 3, 4, and so on.

3. Event Occurrence: From the Include patients dropdown list, select the number of occurrences of the event or no occurrences of the event.

Select No to define patients who do NOT have a specific event (in this case, patients without the defined Inclusion Criteria are filtered out of the cohort).

When moving the cursor over each option in the dropdown list, an explanation and example is displayed.
4. Set the **Age at event** values by entering the required age values in the age field(s).

5. **Filters**: To add additional filters, click **+ Event filter** to further narrow the event's definition (for details on how to define events, see **Defining Event Properties**).

6. To define additional Inclusion Criteria, click **+ Inclusion Criteria**.

7. **Demographic Criteria**: After defining one or more Inclusion Criteria, define Demographic Criteria (optional), if required. To define Demographic Criteria, click **+ Demographic Filter**.

8. To delete a defined Inclusion Criteria, click the x icon on the right.

9. After completing the addition and filtering of inclusion criteria, click the **Save** icon in the upper right hand corner to save your changes.

10. To calculate the cohort, click **Calculate Cohort** at the bottom of the screen (for more information, see **Calculating the Cohort**).

11. To perform additional actions, click the **Action** icon and select the required action from the **Action** dropdown list (see **Performing Additional Actions on Queries**).

12. To continue to the next step in the query definition process, click **Next** or click the **Time-Related Events** tab.

Continue to one or more of the following steps:

- **Step 3: Adding and Defining a Time-Related Event**
- **Defining the Output Columns**
3.3. Defining Events and Properties in the Query Tool

This section provides detailed descriptions of the features and functionality provided in the Query Tool to search for and define Events, Properties, and their Values. Although these features can be used intuitively, this section describes in detail the full functionality of each of the following features:

- Defining an Event Using the Search Engine
- Defining an Event with the Value Browser
- Defining an Event Manually
- Defining the Occurrence of an Event
- Defining Event Properties
- Performing Additional Actions when Creating Queries
  - Defining a Query with No Reference Event
  - Redefining a Reference Event
  - Copying an Event Definition

3.3.1. Defining an Event Using the Search Engine

The Search Engine can be used to assist and simplify the definition process of an event. The typed string runs in MDClone’s database and displays its matches. You can search for Events, Properties, or Values. By default, the scope of the search is for Values. The Search function is displayed slightly differently in each screen where a search is made.

To define an event using the Search Engine:

1. In the Event screen (for example, when defining the Reference Event or Inclusion Criteria), click the Search values properties or events search box.

2. Type a string of at least two characters. While typing the string, a lightbox opens and the string is typed in it. The results are displayed in the Search Results screen.
3. Select the required **Event**, **Property**, or **Value** from the **Search Results** screen.

The following information and options are available:

- **Scope of the search**: Select the scope of the search results that you require, either from the **Events**, **Properties**, or **Values** (default) tab. All results are displayed in the **Values** tab. Tabs that do not have search results are disabled.

A tooltip indicates if no results were found in Events or in Properties. For example:

- **Show My Events** (default): If selected, the search is also performed in the user’s private events (in **Private Loaded Events** and **Combined Events**) in addition to the events in the database that were mapped by the admin. If you do not want search results from your Private Events deselect this checkbox (for details about private events, see **Combined Event Overview** and **Creating a Private Loaded Event**).

- **Did you mean**: Provides suggestions, for example, if the searched term is misspelled.

- **Search Results list**: Select the required values from the Search Results list.

The following information is displayed for each result in the **Values** tab:

- **Group**: Shows the Display Group icon, if the events (including Private Events) were mapped to a Display Group by the admin.
- **Event Name**: Name of Event in which the search results were found.
- **Property Name**: Name of the Property in which search results were found.
○ **Found Values**: The number of values in which search results were found. Clicking a row in the list opens the Value Browser (see Defining an Event with the Value Browser).

4. To see details about the event and its property, move the cursor over a row in the result list. The Details pane (see Details Pane) is displayed on the right, with some or all of the available information:
   - **Group**: The name and icon of the event’s Display Group.
   - **Event Description**: A description of the event as provided by the admin. If more information was linked to the event, click the More details link to open a PDF file with more information.
   - **Property Description**: A description of the property, provided by the admin.
   - **The available data is since**: The earliest year from which data (with at least 500 rows) in the property is available.
   - **% Populated**: The percentage to which this property is populated in the database.
   - **Data Type**: One of the following three data types (and icon) is shown: Numeric, String, or Date.
   - **Data Sample**: A sample of the data.

5. **Value Browser**: In the Values tab, in the list of Values, click the row with the Found Values link to open the Value Browser and see the Values found in this property.

The Value Browser appears, where the user can browse through the list of values to filter and select the required values. The selected events or properties used in the search are displayed beneath the Search box.

Note that if there are no filters available for a search result, the Filters area (in the left pane) is not displayed.
For a detailed description on how to select values with the Value Browser, see Defining an Event with the Value Browser.

6. After selecting required values, click **Apply Selection**. The user is returned to the Reference Event screen (or the screen where the Search was performed). The selected events or values are now displayed in the screen where the search was performed.

7. To save your changes, click the **Save** icon.

8. To continue, click **Next** or select the next step in the Query Wizard.

### 3.3.2. Defining an Event with the Value Browser

When the user defines a Reference Event or Inclusion Criteria either manually or using the Search function, the search results screen is displayed showing all instances (either in the Event, Properties, and/or Values tab) where the search string is found in the database. The displayed values are generally values in the Categorizations that were uploaded to this environment during the data configuration and mapping stage. The categorization functionality enables the user to automatically classify the search results into categories. This functionality can help the user to narrow down the results list, by selecting the categories that are most relevant to his search.

**Note**: Not all events are mapped into categories. Some events are mapped into standard categories, for example: Diagnosis/Conditions are mapped into ICD, which stands for International Statistical Classification of Disease and health related problems, as well as to additional categories based on the organization policy (for example, Body part, Severity, laterality, etc.).

Medications are mapped into ATC which stands for Anatomical Therapeutic Chemical Classification System, as well as to additional categories based on the organization policy (for example, Form, Dosage, etc.).

In the following example, the search term entered in the Search box was **atherosclerosis**. The Search Results display the following found values in the Values tab. The **Found Values** column shows the number of times the relevant codes were found in the database.
To define values using the Value Browser:

1. Enter a term in the Search function or define an event manually, as described in Defining an Event Using the Search Engine or Defining an Event Manually.

2. Click the Values tab in the Search Results screen, to display the following information for the Found Values: Group, Event Name, Property Name, and Found Values. Select the row in which you want to view the details of the Found Values. The Details pane is displayed showing more information for the Event and Property where the string was found (see Details Pane).

3. To filter out private events, if they are not required, deselect the Show My Events checkbox.

4. To open the Value Browser and view the list of the Found Values, click the required row.

The Value Browser opens, where you can search for the required values from the categorizations that were uploaded to the database during the data mapping and configuration process. The Value Browser displays all values where the search string was found in the default or in the selected Category.

The following figure shows an example of a partial list of values displayed in the Value Browser after entering a search string for **atherosclerosis**.

5. **Filters**: Use the filters in the Filters area in the left pane to help narrow and filter the values. The Filters list displays the list of the categorization values.
**Note** that if there are no available filters, the left pane with the Filter list is not displayed. Only the right hand pane is displayed.

- In the **Filters** list, select required values from the Categorizations in order to narrow the list of values in the right pane. Open each Category header (using the > arrow) to display the values in that Category column. Select the required values in that column.

When values are selected in the **Filters** list, the List of Values in the **right hand pane** is filtered, and displays only the values that contain the selections.

- **Use the Search** function in the **Filters** list to search and filter for specific codes or values.

6. In the **right pane** of the Value Browser, select the checkboxes of the required values in the Values List. The following features can be used to help make the selections:

- **Select the Category Level**: The values in the default Category Level (usually the Description or Designation column in the Categorization file) are displayed in the Values List. This is the default category that was configured during the data configuration stage.
If required, select a different **Category Level** from the dropdown list. When defining a generic query (that the user might want to re-run several times), the user might prefer to select values from a higher category level instead of selecting specific codes from the default level. After selecting a different category, the values in the Values List change, depending on the selected Category Level (the selected column in the Categorization file).

- **To view an Additional Category Level:** To display the values of an additional Category Level, click the selection icon in the upper right hand corner. A dropdown menu lists the additional **Category Levels**. Select an additional Category Level to be displayed in the Values List. This view can support the user in the selection of values, by verifying that the selected codes are indeed the required values.

For example, if the default Category Level is **Original Name**, you can add an additional Category Level of the **ICD10 Code**. The Value List then displays values found in both Category Levels (only where the values are found in both categories). In the following example, both the **Original Name** and the **ICD10 Codes** are displayed:
Sort the Values by Usage Frequency (default): Select values that are used more frequently by using the Usage Frequency sorting option that displays the range of frequency that the specific value appears in the organization’s data (from high to low frequency).

A tooltip describes the frequency range (for example, >1K indicates a frequency of 1000 to 10K).

Note that the frequency is calculated for the entire organization regardless of the specific user’s cohort.

Sort the Value List Alphabetically: Sort the Value List alphabetically (A to Z) to help search for specific terms.

Advanced Search: If required, use the Advanced Search option to the right of the Search box, to further refine your search. Click Advanced to display the advanced search options.
The **Advanced Search** screen appears and displays the following advanced search options:

- **All of these words**: Type words to search for values that include all words entered in the text box.
- **Any of these words**: Type words to search for values that contain even one of the words entered.
- **None of these words**: Type words to search for values that do **Not** include the typed words.
- **This exact word/phrase**: Type words to search for values that include only the **exact** typed word or phrase.

Enter the words for the **Advanced Search** options, and click **Search**. The Search Results List now displays only the values that meet the conditions entered in the Advanced Search feature.

8. **Select the checkboxes of each required value**: After searching, filtering, and sorting values, either select **all values** in the Values List or select the **specific** values that you require.

   **Important**! It is not enough to search or apply the filters in order to select the values. After applying the filters, the **required checkboxes in the Value List must be checked** before clicking **Apply Selection**. If values are not selected in the Value List, they will not be saved!

9. **After selecting the values in the checkboxes**, to view the selected values, click **Show** at the bottom of the screen. For example, the following screen opens showing the list of selected values.
Review the list and delete values that are not required.

10. If required, add new values at any stage by performing a new search and repeating the steps to add and select additional values.

11. After selecting all required values, click **Apply Selection**. The screen where the search was made appears (for example, the Reference Event screen) and displays the list of selected values. The following example shows the Reference Event screen after selecting the coronary events using the Value Browser. You can make additional searches, if necessary, from the Search box displayed over the selected values.

If an additional search is made, the new values are added to the original selected values. The **Show values** option now displays and also highlights the new values.
12. After selecting all new required values, click **Apply Selection**.
   The user is returned to the screen from which the search was made and the list of selected values is displayed.

13. To save your settings, click the **Save** icon in the upper right corner of the screen.

   ![Note that you cannot calculate the cohort size at this stage without saving the selection!](image)

14. To calculate the cohort size, click **Calculate Cohort** at the bottom of the screen (for more information, see [Calculating the Cohort](#)).

   The calculated cohort size and date of the calculation is displayed.

15. To continue, click **Next** or select the next step in the Query Wizard.

### 3.3.3. Defining an Event Manually

Defining a Reference Event or Inclusion Criteria manually enables the user to define an event and its filters one by one. Manual definition can be useful when the user wants to change a setting in an event that is already defined and saved, rather than redefining the entire event.

**To manually define an Event:**

1. In the Event screen (for example in the Reference Event or Inclusion Criteria screen), click **Define Manually**.

   ![Narrow your cohort by defining additional inclusion criteria. These criteria are not related in time to the Reference Event.](image)

   ![Note that the options for selecting and filtering an event are displayed slightly differently in each step of the wizard.](image)

2. Define the following by either typing a value, searching, or selecting a value from the dropdown menus:
● Select the **Event Occurrence** from the Get dropdown list (see *Defining the Occurrence of an Event*).

● Define an **Event** manually by either typing a string or choosing an option from the Events dropdown list. The events in the dropdown list are divided into Display Groups, such as Medications, Diagnostics, and Laboratory.

Move the cursor over an event in the list to see more information displayed in a **Details** pane to the right (see *Details Pane*).

**Add a Combined Event from the dropdown list:** If required, define a Combined (And or Or) Event by clicking the **+ New Combined Event** option at the bottom of the event dropdown list (this option is only available when defining a Reference Event).

The Combined Event screen appears.
Define the combined event as described in Creating a Combined (AND) Event or Defining a Combined (OR) Event. After clicking Save and Generate, you are returned to the original screen within the Query tool where you can continue with the query creation process.

- Define the **Age at event**.
- Add an **Event filter** to narrow the event’s definition by defining its properties.
- If necessary, search for a property using the Search Engine and the Value Browser. For details, see Defining an Event Using the Search Engine and Defining an Event with the Value Browser.
- Define a **condition** field after selecting a Property (for details, see Defining Event Properties).

  ![Define the combined event as described in Creating a Combined (AND) Event or Defining a Combined (OR) Event.](image)

  - Define the **Age at event**.
  - Add an **Event filter** to narrow the event’s definition by defining its properties.
  - If necessary, search for a property using the Search Engine and the Value Browser. For details, see Defining an Event Using the Search Engine and Defining an Event with the Value Browser.
  - Define a **condition** field after selecting a Property (for details, see Defining Event Properties).

  Note that the displayed fields and menus depend on one of the following property types: **Numeric**, **String**, or **Date**.

3. To delete a defined filter from an event, click the X icon to the right of the filter’s section.
4. To continue, click Next or select the next step from the wizard.

### 3.3.4. Defining the Occurrence of an Event

The **Get** dropdown menu is used when defining which occurrence of an event is required for the query (for example, when defining a Reference Event, a Time-Related Event, or a Combined Event).

The **Get** dropdown menu enables the user to select the most appropriate occurrence of the event. The required occurrence might be the first, second, or last event that occurred within...
the time frame. However, it is also possible to select the first, second (or Nth) most recent event, by selecting the most recent checkbox.

Note that available options differ depending on the step that is being performed in the query creation process.

- Note that when defining a Reference Event, it is only possible to select a Single Occurrence of the event (first, second, most recent, latest, etc.).
- However, when defining a Time-Related Event, it is possible to select either a Single Occurrence Event or a Multiple Occurrence Event. For a description of Multi Occurrence Events, see Defining a Multiple Occurrence Event.

A tooltip is provided describing each of the Get dropdown menu options. The following figures show examples of the tooltips displayed for the Get occurrence options.

- First: The selected event will be the first earliest occurrence after or before the Reference Date. For example:
● **Second** (or other n value): The selected event will be the second (or third, etc.) earliest occurrence after or before the Reference Date. For example:

![Diagram](image)

Example
Select the second diagnosis event that diagnosis date is after 2013.

allowed date range

2013

● **Max**: The selected event will be the event with the maximum numeric value of the property that will be selected next. Its occurrence will be after or before the Reference Date. For example:

![Diagram](image)

Example
Select the event with the maximum glucose value.

allowed date range

Glucose 82 155 120

● **Min**: The selected event will be the event with the minimum numeric value of the property that will be selected next. Its occurrence will be after or before the Reference Date. For example:

![Diagram](image)

Example
Select the event with the minimum glucose value after admission date (Reference Date).
- **Median**: The selected event will be the event with the median numeric value of the property you will select next. This event’s occurrence will be after or before the Reference Date. For example:

![Median Example]

- **Count**: Count of the event’s occurrences after or before the Reference Date. For example:

![Count Example]

- **Aggregation functions**: When defining Output Definitions, various aggregation functions (count, average, sum, etc.) can be applied to the numeric property that will be selected next. This function can only be applied to a numeric property. This event’s occurrence will be after or before the Reference Date. For example:

![Aggregation Functions Example]
If **Aggregation functions** is selected, the **Output Columns** tab (when defining Output Definitions) provides the option to select which aggregated functions to display.

Note that **Aggregation functions** can be used to define multiple values of a Property rather than defining them one by one. For example, if a user requires the max, min, and median value for a lab result (for example, hemoglobin), the **Aggregation function** enables the user to select all of these values in the **Output Columns** tab instead of defining a max, min, and median value individually for the property. For example:

The example shows the output columns that can be selected when **Aggregation functions** is selected for a numeric property. The user can select all the functions in one location rather than defining each function individually.

### 3.3.5. Defining Event Properties

When defining Event Properties in the Query Tool, various options can be selected from displayed dropdown menus. The displayed options depend on whether the Property type is **Numeric**, **String**, or **Date**. The menu options can be selected intuitively. This chapter provides a summary of the available options.
To define event properties:

1. When defining an Event in the Query Tool, select the required Property of the Event. The Property type can be either a **Numeric**, **String**, or **Date**.

2. **To define a Numeric Property:** After selecting a numeric property from the dropdown list (for example, age at event) select the required **condition** from the dropdown list.

   - Please select condition
     - is less than
     - is less than or equal to
     - is greater than
     - is greater than or equal to
     - is between
     - is not between
     - is equal to
     - is not equal to
     - is empty
     - is not empty

   For all fields (excluding **is between** and **is not between**), type a number (**value**).

   For **is between** and for **is not between**, type a range of two values.

3. **To define a Date Property:** After selecting a date property from the dropdown list (for example, medication start date), select the required **time function** from the next dropdown list. The default option is **year**. Then choose a **condition**. The displayed **condition** options depend on the selected time function.

   For **Date and Hour**, select one of the following conditions:
After selecting a condition for the **date**, define the date using the Calendar feature. If selecting **is between** or **is not between**, define two dates to set a date range.

If the selected **date function** is **hour**: Define the hour and minutes.

For **weekday, week, month, quarter, and season**, select one of the following **conditions**.

Then choose a **value** from the dropdown list (for example, for **weekday**).
For **Year**, select one of the following conditions and then define the year **value**. If selecting **is between** or **is not between**, define two year values to set a range of years.

4. To **define a Text (String) Property**: After selecting a **String** property from the dropdown list (for example, prescribed medication), select the **condition**. The default option is: **is any of the following**.
   It is possible to define multiple strings with an OR relation between them.
3.4. Performing Additional Actions when Creating Queries

The user can perform various additional actions during the query creation process. The following Action icons can be clicked to display the list of additional actions.

- Additional Actions on Queries: During the five steps of creating a query, the upper right hand corner of each screen displays the following icons: Update, Save, and
Action icon. Click the Action icon to display the list of actions that can be performed on the query. For example:

The available icons and actions depend on the specific stage of the query creation process. For more information on actions that can be performed on queries, see Performing Additional Actions on Queries.

- Additional Actions when Defining Events: An additional Redefine icon and an Action icon is displayed on the upper right corner of the Event area when defining Events (for example, Reference Events, Inclusion Criteria, or Time-Related Events). The available actions depend on the specific stage of the query creation process.

The following additional actions are described in this section:
  - Defining a Query with No Reference Event
  - Redefining a Reference Event
  - Copying and Pasting an Event Definition
3.4.1. Defining a Query with No Reference Event

To define a query with No Reference Event:

1. Click the Action icon to the right of the event area and select No Reference Event.

   ![Action Icon](image)

If No Reference Event is selected, the following is displayed and the Birth Date is used as the Reference Event.

![Birth Date as Reference](image)

2. If a Reference Event was already defined and the query was saved, a confirmation message appears. Click Yes to define the query with No Reference Event.

3.4.2. Redefining a Reference Event

The user can change or redefine the Reference Event. This action does not apply a filter to the current reference event, but rather chooses a different event and properties.

To redefine the Reference Event:

1. Open the Query Tool to the required Event definition.
2. Click one of the following options.
   - Click the Redefine icon to the right of the event’s row.
   - Click the action icon to the right of the event’s row and select Redefine.
3. Redefine the Reference Event using the Search Engine or Manually.
4. Complete the steps in defining the event as explained in Defining the Reference Event.
3.4.3. Copying and Pasting an Event Definition

The user can copy an Event definition if it is useful for more than one event (within the same query or between different queries).

To copy an Event Definition:

1. Select the Event whose definition you want to copy and click the Action icon in the upper right side of the Event screen.
2. Click Copy.
   The Event definition is copied to the clipboard and can now be pasted.
   Select the Event whose definition you want to create by pasting the copied definition. Click the Action icon in the upper right side of the Event screen.
3. Click Paste.
   The first Event definition is pasted as a definition into the second Event.

   Note that the Time definition and the Output Columns are not copied.

3.5. Calculating the Cohort

When the user calculates the cohort, the cohort size is based on the cohort definition at the time of calculation. Choosing Calculate Cohort, recalculates the current cohort based on the changes in the cohort definition. The cohort size can only be calculated at specific steps in the query creation process. It is defined and can be updated during Step 1 (Step 1: Defining a Reference Event), Step 2 (Step 2: Defining Additional Inclusion Criteria), and Step 5 (Step 5: Finalizing the Cohort and Output) of the query creation.

   It cannot be recalculated during Step 3 (Step 3: Adding and Defining a Time-Related Event) and Step 4 (Step 4: Defining Patient Demographic Columns) when defining the Output Definition.

   Note: In some cases, it is recommended to make a duplication of the query so the original cohort is not changed. This enables the user to work on a query with the same definitions.

After the cohort is calculated, the footer displays the cohort size and the last date of calculation. The cohort size is displayed differently, depending on the user’s permissions (for a description of the different user permissions, see Appendix B - User Permissions).

- Users with permission to access original data can view the precise cohort size:
Users with permission to access only synthetic data (and who do not have an Approval Key) can only view an approximate cohort size:

OR

The information icon (i) might display the following:

To calculate the cohort:

1. To calculate the cohort during the query creation process, do one of the following:
   - For a query whose cohort has never yet been calculated, click Calculate Cohort at the bottom of the screen. The cohort size is displayed with the calculation date.
   - For a query that has been previously calculated and some of its definitions have changed, a Recalculate Cohort icon is available. Click the Recalculate icon.

A confirmation message is displayed.
To continue, click **Yes, recalculate**.

The **Calculate** icon indicates that the cohort is being calculated. When the calculation is completed, the new cohort size is displayed with the calculation date.

2. To recalculate or abort the cohort while it is still being calculated, click **Recalculate Cohort**.

   The current calculation is aborted.

3. **Information icon**: To view user limitations related to the cohort size, move the cursor over the Information icon. The following different kinds of information might be displayed:

   - If there are no limitations on events (the allowed cohort was not defined):

     ![COHORT LIMITATIONS](image)

     - Max Size: 10,000 Patients
     - Date Range: Unlimited to Unlimited
     - For more details: My Permission page

   - If there is a date range limitation on the cohort:

     ![COHORT LIMITATIONS](image)

     - Max Size: 100,000 Patients
     - Date Range: Oct 1, 1898 to Jan 9, 2301
     - Only events that occurred during this timeframe are queried
     - For more details: My Permission page

   - If the cohort size is larger than the allowed cohort size:

     ![COHORT LIMITATIONS](image)

     - The query's cohort is bigger than your allowed cohort/Max size. The cohort will be reduced to the allowed size in the output definition.
● The cohort size is not up to date and does not reflect the latest data updates:

```
<table>
<thead>
<tr>
<th>COHORT LIMITATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>▶️ The cohort size does not reflect the recent changes.</td>
</tr>
<tr>
<td>Max Size: 100,000 Patients</td>
</tr>
<tr>
<td>Date Range: Oct 1, 1898 to Jan 9, 2301</td>
</tr>
<tr>
<td>Only events that occurred during this time-frame are queried</td>
</tr>
<tr>
<td>For more details: My Permission page [ ]</td>
</tr>
</tbody>
</table>
```

● If the query is a shared query, the query owner’s permissions are used:

```
<table>
<thead>
<tr>
<th>COHORT LIMITATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is a shared query.</td>
</tr>
<tr>
<td>Query shared by: Researcher One</td>
</tr>
<tr>
<td>The query owner’s permissions will be used:</td>
</tr>
<tr>
<td>Max Size: 10,000 Patients</td>
</tr>
<tr>
<td>Date Range: Oct 31, 1900 to Dec 1, 2330</td>
</tr>
<tr>
<td>Only events that occurred during this time-frame are queried</td>
</tr>
</tbody>
</table>
```
4. Defining the Output Definition

Steps 3 and 4 of the Output Definition consist of defining the data in the Time Related Events (Step 3) and in Demographics (step 4). This is the data that will be generated in the output file.

<table>
<thead>
<tr>
<th>COHORT DEFINITION</th>
<th>OUTPUT DEFINITION</th>
<th>SUMMARY REVIEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Reference Event</td>
<td>3 Time-Related Events</td>
<td>5 Finalize Cohort &amp; Output</td>
</tr>
<tr>
<td>2 Additional Inclusion</td>
<td>4 Demographics</td>
<td></td>
</tr>
</tbody>
</table>

Defining the Patient Cohort: In the first two steps of the query, the user defines the Patient Cohort, which is the group of patients whose defining quality is one or more Reference Events (and possibly additional Inclusion Criteria events) that have occurred.

Defining the Output Definition: In the next two steps, the user defines the query’s output. For the defined patient cohort, the output consists of the patients’ events and data to extract into the output file. These events are generally time-related events, meaning that they are defined in time in relation to the Reference Event. Therefore, for the defined patient cohort (for example, a group of patients diagnosed with coronary atherosclerosis), the output will include the Time-Related Event data (for example, readmission to the hospital after the Reference Event).

The following figure shows the relation over time between Time-Related Events and the Reference Event.

Note that while the Reference Event is used to define the patient cohort, the Time-Related Events do not limit or filter the cohort. Rather, they define the Events that have occurred in time to the patients who are included in the defined cohort. They add information related to the previously defined cohort.
4.1. Step 3: Adding and Defining a Time-Related Event

Step 3 in the Output Definition process is defining a Time-Related Event. The user defines an event (or multiple events) that occurred in time in relation to the patient’s Reference Event. For example, these events might be earlier or later diagnoses, administered medications, or various procedures performed on the patient.

The definition of every Time-Related Event consists of an event definition, its timing, and choosing its output columns.

⚠ Important! After defining each Time-Related Event in the Event Definition tab, make sure you define the Output Columns for each defined event!

To add and define a Time-Related Event:

1. After defining the Reference Event and Inclusion Criteria, click Next or select the Time-Related Events tab in the wizard.

The Time-Related Events screen appears. This screen displays the previously defined Reference Event. After defining the Reference Event in Step 1, a Time-Related Event is automatically generated for the Reference Event and is displayed first in the event list.

For example, a Reference Event was defined in Step 1.

- This event is displayed as the first Time-Related Event in the Event List in the left pane (displayed with a red dot).
- The Event Properties are displayed in the Event Definition tab in the right pane.
Note that if the Reference Event is defined with a Combined (OR) Event type, the events are listed one after the other at the top of the list. If no Reference Event was defined, the Event Definition area is empty.

The Time-Related Events screen consists of the following:

- **Left Pane:** Displays the list of the Time-Related Events added by the user. The first event in the list is the Reference Event.
- **Right Pane:** Provides the following tabs:
  - **Event Definition tab:** In this tab, the user defines each new Time-Related event.
  - **Output Columns tab:** In this tab, the user defines which columns to extract into the output file.
  - **Data Preview tab:** Provides a data preview of a selected event.

2. To add a new Time-Related Event, click **+ New Event** in the upper part of the left pane. The Search engine appears.

3. Define a new Time-Related Event (**Event**, **Property**, or **Value**) that will be related in time to the Reference Event, by using the **Search** function (see Defining an Event Using the Search Engine) or by a manual search (see Defining an Event Manually). If required, use the Value Browser (see Defining an Event with the Value Browser). In the following example, a Procedures event was added. The screen now adds a **Timing Definition** area to define how the Procedures event is related in time to the Reference Event.
4. In the upper part of the **Time Related Events** tab, type a name for the Time Related Event in the event field, for example, **Admissions**. The left hand pane now displays the new event name, **Admissions**, under the Reference Event or under previously defined events.

5. In the right hand pane, in the **Get** field, select the Event from the dropdown menu. Use the **Event Filter** to further define the Event properties and values (for details, see [Defining Event Properties](#)).

6. Continue to define each new **Time-Related Event** and set the timing definition, by performing the following steps, as described in the following sections:
   - **Setting the Timing Definition of the Time-Related Event:** Follow these steps to configure the timing definition of the new Time-Related Event in relation to the Reference Event.
   - **Defining a Time-Related Event with Slider:** Use the slider to more precisely define the timing definition of the new Time-Related Event.
   - **Defining the Output Columns:** For each Time-Related Event (including the Reference Event), define the columns to be generated in the Output file.
   - **Performing Additional Actions on Time-Related Events:** Perform additional actions, as required, on the Time-related Events.
4.1.1. Setting the Timing Definition of the Time-Related Event

Timing Definition: In the lower section of the Event Definition tab, define the timing of each Time-Related Event in relation to the Reference Event (or to a Secondary Reference Event). The Timing Definition of each Time-Related Event includes the following steps:

- **Get the Occurrence of the Time-Related Event** (Single or Multiple, first, second, nearest, multi-occurrence, etc.).
- **Define the Time Relation of the Time-Related Event to the Reference Event** (for example, it occurred before, after, at & after, at the same time, etc.).
- **Select the Reference Event** (select either the previously defined Reference Event or a Secondary Reference Event).
- **Define the Date Property** of the Reference Event.
- **Use the slider** to define the Time-Related Events more precisely.

To set the timing definition of the Time-Related Event:

1. In the Timing Definition area, Get the Occurrence of the Time-Related Event. From the Get dropdown menu, select the Occurrence of the defined Time-Related Event. The default option is the first Time-Related Event. The following options are available when defining the Occurrence of a Time-Related event.
   - **Single Occurrence**: This option enables you to define a single occurrence of the event.
- **Multiple Occurrences**: This option enables you to define multiple occurrences of an event (see Defining a Multiple Occurrence Event).

A tooltip describes each menu option. Some of the options require additional input. For example:

- **Nearest**: The timing definition in the next dropdown list automatically changes from after to Ever (before, at, or after).
- **Max…, Min…, Median…, Count, Aggregate Functions**: The timing definition requires an additional numeric property of the Time-Related Event, as in the following figure.

For a detailed description of the available options when selecting the event occurrence from the **Get** dropdown list, see Defining the Occurrence of an Event. For a description of the options available for defining a Multiple Occurrence Event, see Defining a Multiple Occurrence Event.

2. **Define the Time Relation to the Reference Event**: From the next dropdown menu, define the relation of the Time-Related Event to the Reference Event. The default option is at & after the Reference Event (or the Secondary Reference Event).
Additional options are **at, after, before, ever, or between two dates** in relation to the Reference Event.

![Timing Definition]

3. Select **one of the following** as the Reference Event:
   - **Select the previously defined Reference Event** from the dropdown menu. The displayed (default) Reference Event is the event that was defined during the defining Reference Event stage (if no Reference Event was defined, the birth date is used as the default Reference Event).

![Timing Definition]

The Reference Event is displayed in the Event List and in the event dropdown menu with a red dot.

- **Select a Secondary Reference Event**: If required, select a Secondary Reference Event from the dropdown menu. Any Time-Related Event from the list of the events in the left pane can be defined as a Secondary Reference Event in this dropdown list.

  Note that it is **not** possible to define a Secondary Event on a previously defined Secondary Event.

  Note also that an event that was previously used as a Secondary Reference Event is indicated in the dropdown menu with a red dot. For example:
4. **Define the date property** of the selected Reference Event. The default date (based on the system settings) is selected by default.

If the Time-Related Event occurs **between two dates** of the Reference Event, an additional dropdown list with a **date property** is provided.

5. After setting the Timing Definition, the timeline displays the timing definition of the Time-Related Event in the following ways:
   - **Before**: If the Time-Related Event is defined as before the Reference Event, the slider is highlighted in the area to the left (**before**) of the Reference Event.
   - **At**: If the Time-Related Event is defined as occurring **at the same time as** the Reference Event, the slider displays a blue dot **at the time** of the Reference event.
   - **At and After**: If the timing definition includes both **at and after** (or **before**), the slider displays both conditions in relation to the Reference Event.
6. To set the timing more precisely using the slider, check the **Set timing on slider** checkbox.

For a detailed description of how to define the timing of the Time-Related Events using the slider, see [Defining a Time-Related Event with Slider](#).

7. If the slider is not required, continue by clicking **Next** and by following the steps in [Defining the Output Columns](#).

For additional actions that can be performed, see [Performing Additional Actions on Time-Related Events](#).

### 4.1.2. Defining a Multiple Occurrence Event

When defining a Time-Related Event, it is possible to select multiple occurrences of that event for all patients in the defined cohort instead of only one occurrence (first, second, etc.) for these patients.

**Note the following!**

- This option is only available for Time-Related Events and is not available for defining a Reference Event.
- This option is available only for a user who has permission to query Original Data or a user who can only query Synthetic Data but has a valid Approval Key.

To define a Multiple Occurrence event:

1. Define a Time-Related event, as described in [Step 3: Adding and Defining a Time-Related Event](#).
2. In the **Timing Definition** area of the Time-Related Event (Step 3), open the **Get Occurrence** dropdown menu.
3. Under **Multiple occurrences**, select a value (up to a configurable maximum number defined by the admin in the System Settings) to define either the **first** x occurrences or the **latest** x occurrences of the defined event. By default, this value is up to **99** occurrences. These occurrences might be before or after (or on) the Reference Event, depending on how you define the Time Relation of the Time-Related Event to the Reference Event (before, after, etc.).

4. After entering the number, click the arrow to apply the setting. The event in the left pane Event List now displays the Multi-Occurrence icon, to indicate that it is a multiple occurrence event.
Note that this icon is displayed in Step 3 even for a user who does not have permissions to define a Multi-Occurrence Event.

The percentage displayed in the Event List for a Multiple Occurrence event represents the percent of patients who have at least one event of a Multi-Occurrence event type (as shown in the tooltip).

Note the following limitations related to defining Multiple Occurrence events:

- If a user who does not have permissions (for example, a user with only Synthetic Mode permissions) tries to define a Multiple Occurrence event, an error message is displayed. However, the system does not delete or change the definition at this stage (Step 3).

- Note that if the user who does not have permissions tries to generate an output file with the multiple occurrence events in Step 5, the output file will not include the multi-occurrence events. The following message will be displayed (in step 5):
A Secondary Event cannot be defined as a Multiple Occurrence event. If a user tries to define a Multi-Occurrence event as a Secondary Event the Secondary Event remains a Single Occurrence event and a copy of the event is added to the Event List. Only the copy is created as a Multi-Occurrence event.

4.1.3. Defining a Time-Related Event with Slider

The slider can be used to define the timing of the Time-Related Event more precisely in relation to the Reference Event. For example, you can define an event as up to 100 days after the Reference Event.
Note that the slider is relevant only if the Time-Related Event is After, Before and Ever in relation to the Reference Event. It cannot be used for the At and Between two dates options.

To define a Time-Related Event using the slider:

1. After defining a Time-Related Event (see Adding and Defining a Time-Related Event and Setting the Timing Definition of the Time-Related Event), check the Set timing on slider checkbox.
2. Choose the date property of the Time-Related Event from the dropdown list.
3. Select the required time unit: Hours, Days (default), Months, or Years.
4. Use the slider to set the timing of the Time-Related Event. Drag one or both flags on the slider to set the time range value (for example, the number of days before the Reference Event).

The number of days (or months, etc.) is displayed above each flag. The time can also be set by typing the number value in the flag text boxes.

Note that the movement of the flags depends on the timing definitions. For example, if the Get Occurrence is defined as before, only the before flags (on the left) can be moved. In this case, the flags cannot be moved to the right (after).
The following example shows that only the before flag (on the left) is moved.

5. The flag can **only** be moved up to a value of **100**. If a value higher than 100 is required, type the required value in the text box above the flag. The allowed values are an integer or a decimal number.

6. **Splitting or Merging the Time Range:** To define more than one time period with the flags, click the **Split** icon on the right or left end of the slider. This splits the time range on the slider into two time periods and two flags are now available. The following example shows a split time range that is defined as a start date that is **35 to 100 days** before the Reference Event.

To merge the timeframe again so only one flag is available, click the **Merge** icon.

7. **Double click** the flag to change its type from to and back again to its initial state. The following figure shows the different available flag states:

8. After defining each Time-Related Event, **make sure** to define the Output Columns for each event, as described in **Defining the Output Columns**. The selected output
columns will then be generated in the output file. After defining the Output Column for each event, it is possible to view a Data Preview of the event (see Viewing a Data Preview).

9. To save your changes, click the Save icon in the upper right hand corner or proceed with the query and the changes are saved automatically.

10. For additional actions that can be performed, see Performing Additional Actions on Time-Related Events.

4.1.4. Defining the Output Columns

After defining each Time-Related Event and configuring the timing definitions, the output columns of each event must be defined. These columns represent the Event Properties and data that will be generated in the final output file.

To define the output columns:

1. To define the output columns for an Event, select the required Event in the left pane and click the Output Columns tab. The Output Columns screen appears.

2. In the Column Prefix field, enter a column prefix.

3. Output Column checkboxes: Select the output columns that you want to generate in the output file. Note that some events are automatically selected. Select or deselect the events that are required.
After selecting the columns, the **Output Columns** tab displays the number of selected columns of all the existing columns. For example, the following figure shows that 4 out of 12 columns were selected.

If no Output Columns were selected an alert is displayed next to the **Output Columns** tab. A tooltip alerts the user to the fact that no output columns were selected.

4. **Output Column Name**: The name of the output column is displayed as it will appear in the result file. The Output Columns represent the Event Properties. The default name of the column is the Property Name with the prefix concatenated before it. Edit the name of each output column, if required.

By default, only 8 columns are displayed. Click **Show all** to display the full list of the event output columns.

5. **The Information** icon provides a tooltip with a description of the column’s content if this information is provided by the admin during data configuration. To view the information, move the cursor over the icon.

6. If the output is a date, the **Default Date** icon indicates if it is a default date.

7. Under **Date Function/Category**, select how you want a Date or Category to be generated in the output file.

   - **Date**: If the column content is a date, the default output is **Date**.
     
     The following figure shows the options that are available from the dropdown list. For example, it is possible to display the number of days, weeks, or months that the Time-Related Event is before or after the Reference Event.
6. The **Property Name** is displayed and is editable. This is the name of the Event Property as it was mapped in the system by the super admin.

7. To duplicate the output column, click the **Duplicate** icon . The duplicated column is added under the original column with a *copy* suffix added to the new column.

8. Change or edit the duplicated column as required.

9. To save your settings, click the **Save** icon.

10. After defining the Output Columns for the Reference Event and for each of the defined Time-Related Events, click **Next** or continue to one of the following steps:
    - [Step 4: Defining Patient Demographic Columns](#)
    - [Performing Additional Actions on Time-Related Events](#)
    - [Step 5: Finalizing Cohort and Output](#)
4.1.5. Viewing a Data Preview

It is possible to view a preview of the data defined in Step 3 prior to completing the five steps of the query. This enables users to save time and to verify that the data defined in the first four steps is satisfactory for their needs prior to generating the output file in Step 5. The user can make the necessary changes at this stage before proceeding to Step 5.

**Note** the following about the Data Preview:

- A Note explains that it is based on a **data sample** (of 1000 patients) and does not represent the data in the entire cohort. The data sample can display up to a maximum of 100 patients (represented in rows in the table).

- The **Data Preview** tab is only displayed for users who have permission to view Original Data or users of Synthetic Data who have an Approval Key.

- It is only displayed for Single Occurrence events. If the Data Preview was selected for a Multi-Occurrence event, an alert is displayed explaining that it will only display data for a single occurrence of the event (first or latest, depending on the selected option).

To view a Data Preview of an event:

1. In Step 3 of defining a query, in the Event List (left hand pane), select the event whose data you want to preview and click the **Data Preview** tab.
Each **Event Property** is displayed in one column of the table. The displayed Properties are those Properties (without the prefix) that were selected when defining the **Output Columns**. Additional settings defined in the **Output Columns** tab (for example, **Date Function** or **Category**) are also displayed in the Data Preview.

The following information is displayed in the Data Preview table:

- **Property Name**: The first row in the table displays the Property names for this event.

- **Hide/Show Property Description**: The row beneath the Property name displays a description, if it was added by the Admin during the process of environment configuration. This description can be shown or hidden by checking or unchecking the **Show Properties Description** checkbox.

- **Percent Populated**: The next row displays the percentage that the Property is populated within the data sample. This is only a percentage of the data sample and does not represent the percentage populated in the entire cohort.
### Categorical Data:
In columns that display Properties with Categorical data, the row in the table below the Percent Populated row displays only the five most frequent categorical values in the data sample (but not in the entire cohort).

<table>
<thead>
<tr>
<th>DRG Code Desc</th>
<th>DRGMortality</th>
<th>AGE EVENT</th>
<th>DRG Code</th>
<th>DRG Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes</td>
<td>Empty</td>
<td>59%</td>
<td>294</td>
<td>13%</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>Empty</td>
<td>15%</td>
<td>4293</td>
<td>8%</td>
</tr>
<tr>
<td>Coronary</td>
<td>Empty</td>
<td>10%</td>
<td>174</td>
<td>5%</td>
</tr>
<tr>
<td>Septicemia</td>
<td>Empty</td>
<td>6%</td>
<td>107</td>
<td>4%</td>
</tr>
<tr>
<td>Pancreas Transplants</td>
<td>Empty</td>
<td>0%</td>
<td>871</td>
<td>4%</td>
</tr>
</tbody>
</table>

### Numerical Data:
In columns that display Properties with Numerical data, the row below the Percent Populated row displays a Histogram of the data.
Moving the cursor over each bar in the Histogram displays the value represented by the bar. It displays the Percent Populated value and also the actual value (or range of values represented by each bar).

- **Data Values:** The following rows in the Data Preview table display values for each property taken from the data sample (display is up to a maximum of 100 rows). They do **not** display all values that will be generated in the Query Output.
4.2. Step 4: Defining Patient Demographic Columns

Step 4 in the query building process is defining the Demographic output columns. These columns are properties that contain the demographic static data of the patients in the cohort.

To define the demographic columns:

1. Click the **Demographics** tab in the wizard. The following **Demographic Columns** screen appears.

![Demographic Columns Screen](image)

The columns with the demographic static data of the patients in the cohort are displayed in a table with the following information and options:

- **Checkboxes**: Select the checkboxes of the columns with demographic data that you want to include in the extracted file.
- **Column Names** are displayed as they will appear in the generated output file (as originally mapped by the admin). Edit the column names, if required.
- **Information icon**: Displays a description of the column (Property) in a tooltip (as provided by the admin during the data configuration stage).
- **Default Date icon**: Indicates which Date property is the default date.
- **Date Function/Category**: It is possible, if required, to select how a Date or Category output is displayed.

If the column content is a **date**, the default output is **Date**.

The following figure shows options that are available from the dropdown list. For example, it is possible to display the number of days the Event is distant from the Reference Event.
If the column content is from a Categorization, the default Category Level (as configured by the admin) is used. A different Category Level can be selected from the dropdown list.

- **Type icon**: The Property type (Numeric, String, or Date) is indicated with one of the Type icons.
- **Property Name**: The name of the property as it exists in the database.
- **Duplicate Property icon**: Click the Duplicate icon to duplicate a property to be edited or changed as required. The duplicated property is added to the column list with the term copy added after the name (for example, birth date copy).

2. To save your settings, click the Save icon or click the Action icon in the upper part of the screen and click Save.

3. To continue to the next step in building a query, click Next or select the Finalize Cohort and Output tab in the wizard. Continue with the steps described in Step 5: Finalizing the Cohort and Output.
4.3. Performing Additional Actions on Time-Related Events

Additional information and actions are available when defining Time-Related Events. The following figure shows an example of the Event List in the left hand pane of the Time-Related Events screen.

The following features and functionality are available:

- **Hide/Show list**: Click the Hide/Show icon in the upper left corner to hide or show the list of events.

- **Reference Event indication**: The red dot on an Event indicates that it is the Reference Event (or a Secondary Reference Event). The Reference Event is the first event in the Event List. When selecting the Reference Event in the left pane, the event definition is displayed in the right hand pane. The Reference Event definitions are not editable during this step in the process. To redefine a Reference Event, return to the first step in the wizard, where the Reference Event is defined (see Step 1: Defining a Reference Event).

- **Search**: Use the Search function to search for required events from the list of events.

- **Reordering the Event List**: Reorder the events in the list as required, by dragging each event up or down in the list. The order of events in the Events List will be the order that the output is displayed in the final output file.

- **Percentage**: For each Event in the Event list, the list displays the percentage of that event’s occurrence in your defined cohort. The Reference Event is always 100%
populated, because the Reference Event defines the cohort. All other Events might only be populated by some percentage of the cohort. In the following example, only 19% of the patient cohort were readmitted to the hospital.

- **Delete**: To delete an event, click the x icon to the right of the event.

A confirmation message appears. To delete the event, click Yes.

Note that deleting a Secondary Reference Event changes the events referencing it to reference the Reference Event instead.

- **Total Number of Events and Selected Output Columns**: During the process of defining the Time-Related Events and Output Columns, the lower area of the left pane displays the number of defined Events and the number of selected Output Columns. The numbers are updated with each new input.

- **Secondary Reference Event**: While setting the timing definition of a Time-Related Event (see Setting the Timing Definition of the Time-Related Event), it is possible to define it in relation to an Event that is not the Reference Event. For example, you can define a Readmission Event in relation to an earlier Admission Event. The earlier Admission Event is then defined as a Secondary Reference Event.

After creating a Secondary Reference Event, the Event List displays the red dot indication for both the main Reference Event (for example, PTCA) and for the Secondary Reference Event (Admission).

The following figure shows a Readmission Event that is defined in relation to an earlier Admission Event (and not in relation to the Reference Event). The earlier Admission Event is then defined as a Secondary Reference Event.

Note that a Secondary Event cannot be related to a previously defined Secondary Event.
**Additional Actions:** Click the Action icon for additional actions that can be performed on the Events.

- **Duplicate:** Creates a copy of the Time-Related Event and opens it. The new event is added under the original event in the list. The name is the same as the original event with `_copy` added (`<eventname>_copy`).
- **Copy:**Copies the definition of the Time-Related Event without the **Timing Definition** or the configured **Output Columns**. Copying can be performed within a query and also from one query to another query.
- **Paste:** Pastes the definition of an event that was copied from another step or another Time-Related Event (it is pasted without the **Timing Definition** or the configured **Output Columns**).
- **Redefine:** Opens the Search function to enable you to redefine and edit the Time-Related Event.
- **Delete:** Deletes the event. A confirmation message appears. To delete the event click **Yes**.

> Note that deleting a Time-Related Event that is used as a Secondary Reference Event will cause the events referencing it, to reference the Reference Event instead.
5. Summary Review

Step 5 in building a Query consists of finalizing the cohort and the output and generating the output file.

The following figure shows an example of the **Finalize Cohort and Output** screen.

This screen is displayed as a table. Under **Column Name** is the list of all the columns that can be selected for the result file. The user can select which columns to include or exclude from the extracted file, can filter the data, change Category levels, and create new columns with calculated fields.

To complete the query building process and download the output file, perform the steps in the following sections, as required:

- **Step 5: Finalizing Cohort and Output**
- **Adding Calculated Fields**
- **Performing Additional Actions on Queries**
5.1. Step 5: Finalizing the Cohort and Output

After defining the Reference Event, Inclusion Criteria, and configuring the Output Definitions, the user can proceed to the final step of finalizing the cohort and the output and generating the output file.

The Finalize Cohort and Output screen is displayed differently for users who have permission to query Original Data or for users who are only allowed to query Synthetic Data (for more information, see Appendix B - User Permissions).

To finalize the cohort and output:

1. Log in to the MDClone Dashboard. The My Queries screen appears. OR, click Query Tool > My Queries.
2. After configuring the Output Definitions, click Next or select the Finalize Cohort and Output tab in the wizard.
3. In the upper right hand corner a toggle is provided to turn On or Off the option to create Synthetic Data. This toggle is only displayed for users who are allowed to query Original Data (or Synthetic Data users who have an Approval Key). It is not displayed for users of Synthetic Data only.
   Change the Synthetic mode to On or Off as required (for a description of allowed data for different users, see Appendix B - User Permissions).

4. For the Synthetic Data User (only): The following gauge is displayed at the top of the screen:

   This gauge indicates what percentage of rows in the data are censored with one or more censored columns. Censored data is data that might cause a breach in patient privacy (might enable users to identify patients).
   The red area indicates a high percentage of censored data, while the blue and green areas indicate a low percentage of censored data.
   In the following example, there is a high percentage of censored data:
The Synthetic Data User uses the gauge to try to lower the percentage of censored data to a minimum or, if possible, to 0%. For example, this can be achieved by removing the columns with a high percentage of censored data or by changing the categorization of a column.

5. The **Finalize Cohort and Output** screen displays the information defined in the previous steps:

- **Checkbox**: A checkbox is provided next to each Column Name. By default all columns are selected. Select the columns that you want to include in the output.

- **Number of selected columns**: After selecting the required columns, the number of selected columns (out of all configured columns) is displayed above the Column Names. For example, **24 selected of 26**.

- **Column Name**: A list of all the columns that were defined and added in the Output Definition is displayed (see **Defining the Output Columns**).

- **Type icon**: The Property Type is indicated with either the Numeric, String, or Date icon. It is possible to convert a Numeric or Date Type to String (however, a String cannot be converted to a Number or Date).

  Note that **Data Type** cannot be converted for a multi-occurrence event.

If the user moves the cursor over the **Type** icon it is disabled and the following is displayed:

- **Category/Calculated Column**: Indicates if the column is related to one of the following:
  - **Categorization**: If it is related to a categorization, the default Category is displayed. To change the Category, select another category from the dropdown list. The data in the output file will be changed accordingly.
  - **Calculated Column**: Indicates if a calculation was performed on the column (see **Adding Calculated Fields**).

  An **Alert** icon is displayed if there is an error or problem with the calculated field. To edit the Calculated Column, click the **Edit** icon.

- **Information icon**: Displays randomized examples of the column values in a tooltip to show how the data will appear in the output file.
● **Censored (displayed only for Synthetic Data users)**: Displays percentage of censored rows in each column. Only a data type of **String** can be censored. Numeric or Date values cannot be censored. The following is an example of a column with a high rate of censored data.

Censored Data is data that might cause a breach of patient privacy, for example, the patient’s id. In the Synthetic Data output file, a column with the patient ID will be censored and the word **Censored** will replace the data in each cell of the column. For example:

<table>
<thead>
<tr>
<th>internalpatientid</th>
<th>censored</th>
<th>censored</th>
<th>censored</th>
</tr>
</thead>
</table>

● **% Populated**: Displays the percent that the data in the Property is populated in the cohort (what percentage of patients in the cohort have this data value).
  ○ **For Original Data** users (and for users of Synthetic Data with an Approval Key), this column displays the exact percentage.
  ○ **For Synthetic Data** users, this column displays only the percentage range and not the exact percentage.

● **Multiple Occurrence Event Icon** 🗓: Displayed on each row whose event was defined as a multiple occurrence event. This is displayed only for Original Data users (and for users of Synthetic Data with an Approval Key).
● **Additional Actions Icon**: To perform additional actions on a column, click the Action icon on the right side of the row. The following actions are available from the dropdown menu:

- Duplicate
- If Exists Conversion
- Filter

**Note** that Additional Actions are disabled for an event defined as a Multiple Occurrence event. When moving the cursor over the Action icon on the row of a Multi-Occurrence event the following is displayed.

- **Filter icon**: Indicates if a filter was applied on a column. Applying a filter might change the cohort number.

- **Cohort number**: The number of patients in the cohort is displayed at the bottom of the screen with the date the cohort was last calculated. If the data in the screen is filtered or changed in some way, the cohort is automatically recalculated (for more information about the cohort calculation process, see Calculating the Cohort). See also Appendix B - User Permissions.
  - For Original Data users (and users of Synthetic Data with an Approval Key), the exact cohort size is displayed.
    - Cohort: 9,974 | Calculated: Jul 29, 2021 10:39
  - For Synthetic Data users (without an Approval Key), the approximate cohort size is displayed.
    - Cohort: ~9,973 | Calculated: Jul 29, 2021 10:39

OR

   - Cohort: 212-219 | Calculated: Jul 29, 2021 10:51

6. **Filtering**: To filter data in the table, do one of the following:

- Click the Filter button above the Column Name list. The number on the Filter button indicates the number of filters that were applied on the list.
OR

- From the Column Name List, select the column that you want to filter from the dropdown menu, click the Action icon, and select Filter. The Add Filters screen appears.

![Add Filters Screen](image)

The Add Filters screen displays the selected column. It also displays previous filters that were defined in the Column List.

**Note** that the Multi-Occurrence columns are not displayed in the Filter dropdown menu, because filters are disabled for these events.

- Select the condition from the Select condition dropdown menu. The available conditions depend on the data type of the Property (for all available options, see Defining Event Properties) in the selected column (Numeric, Date, or String).

For example, the following condition dropdown menu is displayed when defining a Numeric property:

![Condition Dropdown Menu](image)

- After defining the required filters, click Apply & Close. The filter is applied and the cohort is recalculated.

- Define additional filters by clicking + Column Filter. Then click Apply & Close.

7. **Adding Calculated Fields:** Add and define calculated fields, if they are required. For a detailed description about how to create and add calculated fields, see Adding Calculated Fields.
8. After selecting, filtering, and making changes on the output columns, download the data file (Original Data), by doing one of the following:
   - Downloading Original Data
   - Downloading Synthetic Data

5.1.1. Downloading Original Data

This section describes the downloading of Original Data. It is relevant for users who are allowed to query Original Data and for users allowed to query Synthetic Data but have an Approval Key.

To download Original Data:
1. Log in to the MDClone Dashboard. The My Queries screen opens. Alternatively, click Query Tool > My Queries.
   Create a new query or select an existing query. Click the Finalize Cohort and Output tab in the wizard.
2. In the upper right hand corner, in the Synthetic Mode toggle, make sure that Off is selected.
3. Click Generate Original.
   The following dialog box appears.

   ![Generate Original Dialog Box]

   Note: When there are no occurrences, one empty column will be added to file.

4. Select one of the following kinds of output file(s) that you require.
   - Single Output File
   - Multiple Output Files
5. To view a description of the output files that are generated for each option, click the More Info link.

An Info screen opens that describes the available output file format:

6. **To Generate a Single Output File**: Click the Single Output File option.
   - A tooltip can be accessed that describes the single output file format, as follows:
     - **One** file is generated with all the data.
     - There is one patient per row in the file.
     - Each event occurrence has one column in the file.
     - **Note** that if there are no event occurrences, one empty column is added to the file.
   - To generate the Single Output File, click one of the following options:
     - **Apply and Generate**: The single file is generated. A Success (or Failure) notification is displayed. Click the Download file link in the notification to download the file.
**Apply and Open With**

Click this option to download the file as a CSV file or to open the file with a third party application (if configured by the Admin in this environment). For example:

7. **To generate Multiple Output Files**: Click the Multiple Output Files option.

- From the Events dropdown menu, select the multi-occurrence events that you require, up to the configured maximum value (configured by the Admin in the System Settings).
- Then click Apply (see the following figure).
The Multiple Files format generates the following files:

- One file that contains all single occurrence events with one patient per row.
- One file for each multi-occurrence event, with one row per occurrence.

**Note** the following:

Note: 1. Patient ID will be added to files automatically.
2. When Event is empty (no occurrences), an empty file will be downloaded.

8. After defining the output file format, click **Apply** and **Generate**.

The output files are generated.
A Success (or Failure) notification is displayed.
For example:

To download the files click the **Download files** link in the Success notification.
The system downloads the CSV file(s) as a zip file that includes all the generated files.
5.1.2. Downloading Synthetic Data

This section describes the downloading of Synthetic Data. It is relevant for Original Data users who want to generate Synthetic Data and for users only allowed to query Synthetic Data.

Note that only files with more than 200 rows can be synthesized.

Important! When generating Synthetic Data, the user’s goal is to generate a file with a minimum of censored data. The lower the percentage of censored data, the more the data in the output file is reliable and valid.

To download Synthetic Data:

1. Log in to the MDClone Dashboard. The My Queries screen opens. Alternatively, click Query Tool > My Queries.
2. Create a new query or select an existing query. Click the Finalize Cohort and Output tab in the wizard.
   The Finalize Cohort and Output screen appears.
3. Original Data user: Turn the Synthetic Mode toggle to On.
   Synthetic Data (only) user: There is no On / Off toggle in this screen.
   The Synthetic Mode screen appears and displays the following gauge of censored data that will be generated in the output file.

4. Make required changes in the output columns, if required (see Step 5: Finalizing the Cohort and Output and Adding Calculated Fields).

   Note that Multi-Occurrence columns are unchecked and are not supported in Synthetic Mode. The following tooltip is displayed for each of these columns.

   Note that it is extremely important to attempt to lower the percentage of censored data as much as possible.

   To lower the percentage of censored data, do one or more of the following.

   - Deselect (or delete) highly censored columns.
     For example, the column of internal patient id results in a high rate of censored data. Deselect or delete this column from the list to lower the percentage of censored data.
● **Change a Category level** to another Category with a less granular description. For example, instead of selecting the OriginalName category of a medication, select the Therapeutic/pharmacological subgroup (ATC3) of the medication.

● **Use the IfExists conversion** to convert a highly censored column (see [Performing Data Conversion](#)).

If the changes successfully lowered the percentage of censored data, the gauge is recalculated and indicates the new low percentage of censored data. For example:

![Gauge Image]

5. After completing all changes in **Finalize Cohort and Output** screen, click **Generate Synthetic**.

   ![Confirmation Message]

   Note that if Multi-Occurrence events were included in the Column list, these events become deselected (unchecked) and cannot be included in the generated output files. The following confirmation message is displayed.

6. Click **Yes** to continue.

   The Synthetic Data file is generated. The Synthetic file generation might take time. After generation, a success or failure message is displayed.

   If the file generation is successful, the success message contains links to download the generated files. For example:

   ![Success Message]

7. Download the files. The following files are generated:

   ● **Synthetic Data CSV file**: A CSV file is generated that includes all the configured columns and data that were finalized during Step 5.

   ● **Comparison Report**: A PDF Comparison Report is generated that displays a comparison between the synthesized data and the original data. The comparison file provides validation for the synthetic process. The report includes the following sections:
Introduction that describes the contents and purpose of the Comparison Report. For example:

**Introduction**

The comparison report compares the data of the synthetic file produced by MDClone's synthetic engine, to the data of the corresponding original file. It includes general information about the data, as well as tabular elements that summarize the data in both the original and synthetic files. It also includes a set of measures designated to assess the utility of the synthetic data when compared to the original data.

General information regarding the Synthetic data. For example:

**General Information**

This report was created on Sunday, April 20, 2022, at 13:54. The synthetic data file includes 2000 rows, representing a cohort of that size. The corresponding original data file features a reasonably similar number of rows. Both original and synthetic files include 26 columns, that represent three types of variables:

- 6 categorical variables
- 16 numerical variables
- 4 date variables

Censored data: Information about censored data. For example:

**Censoring**

Some categorical variables are censored to ensure that patient privacy is maintained and no real-life health information is disclosed about an individual or a small group of individuals. The following list depicts the most highly censored categorical variables, and their ratio:

<table>
<thead>
<tr>
<th>Categorical Variable Name</th>
<th>Censored Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>internalpatientid</td>
<td>1.0</td>
</tr>
<tr>
<td>deceased in hospital</td>
<td>0.0</td>
</tr>
<tr>
<td>gendercode</td>
<td>0.0</td>
</tr>
<tr>
<td>cohort reference event-condition code desc</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Variable Distribution: The following table shows the variable distribution. For example:

### Univariate

#### Variable Distribution

The following is a list that shows the coefficients of pairs of variables, for both the original data set and MDClone’s synthetic data:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>gendercode</td>
<td>0.0205</td>
</tr>
<tr>
<td>deceased in hospital</td>
<td>0.0015</td>
</tr>
<tr>
<td>f1</td>
<td>0.0015</td>
</tr>
<tr>
<td>internalpatientid</td>
<td>0.0</td>
</tr>
<tr>
<td>cohort reference event-condition code desc</td>
<td>0.0</td>
</tr>
<tr>
<td>f2</td>
<td>0.0</td>
</tr>
<tr>
<td>cohort agg-age at event.average</td>
<td>0.0</td>
</tr>
<tr>
<td>cohort agg-age at event.count</td>
<td>0.0</td>
</tr>
<tr>
<td>cohort agg-age at event.max</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Multivariate Correlation: Displays a frequency list of pairs of Real data variables and pairs of Synthetic data variables.

### Multivariate Correlation

The following is a frequency list of pairs of Real data variables, and pairs of Synthetic data variables: (Showing 10 pairs with the highest difference)

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Variable Name</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>cohort agg-age at event.quadc</td>
<td>gendercode</td>
<td>0.7828</td>
</tr>
<tr>
<td>cohort count-count</td>
<td>gendercode</td>
<td>0.7828</td>
</tr>
<tr>
<td>decryption-encrypted_date_shift_value</td>
<td>gendercode</td>
<td>0.7828</td>
</tr>
<tr>
<td>cohort agg-age at event.count</td>
<td>gendercode</td>
<td>0.7828</td>
</tr>
<tr>
<td>cohort agg-age at event.sd</td>
<td>gendercode</td>
<td>0.7828</td>
</tr>
<tr>
<td>cohort agg-age at event.slope</td>
<td>gendercode</td>
<td>0.7828</td>
</tr>
<tr>
<td>cohort agg-age at event.quadb</td>
<td>gendercode</td>
<td>0.7828</td>
</tr>
</tbody>
</table>
5.2. Adding Calculated Fields

The user can add various calculations to the data in the output columns by selecting an option from the Calculated Column dropdown menu. Each new calculated column is displayed as a new column in the extracted file. A calculated field enables the user to perform various mathematical operations on the fields in the file. The calculations take place before the synthesis process of the data in the file to preserve the accuracy of the calculations.

The following Calculated Column options are available:

- Custom Expression: For details, see Adding a Custom Expression.
- If exists (true/false): For details, see Performing Data Conversion.
- 2 Column Calculation: For details, see Adding Two-Column Calculation.
- 1 Column Calculation: For details, see Adding One-Column Calculation.

Note that the Multi-Occurrence columns are not displayed in the Calculated Column dropdown menus, because Calculated Fields are disabled for these events.

5.2.1. Adding a Custom Expression

This calculated field enables the user to add a new column with complex mathematical calculations as well as conditional expressions between different columns.

To create a new column with a Custom Expression:

1. Log in to the MDClone Dashboard. The My Queries screen appears.
2. In a new Query or in a previously created Query, select the Finalize Cohort and Output tab to select or display the selected Output Columns.
3. Click the + Calculated Column button above the Column Name list. From the dropdown menu, select Custom Expression. The Custom Expression screen appears.
4. In the **Type a new column name** text box, type a descriptive name for the new customized column.

5. Select the column from the dropdown list on which to perform the operations.

6. From the left hand pane, select the required operations from the list of operations. The selected operations are displayed in the black text area.

Note that for each operation, an example and explanation is displayed when moving the cursor over an operation. For example:
● Use the information provided about the field type during the construction of the expression, to make sure you perform a mathematical operation between two possible fields.
● Make sure that textual fields included in the expression are separated with commas within the scope of the parentheses.
● To perform basic mathematical calculations, select one or more of the following operations: +, -, *, /, ^, (, ), Ln, LOG, LOG10.
● To perform more complicated operations, use the following:
  ● **MIN**: Extracts the field with the lowest value out of the fields entered in this operation.
  ● **MAX**: Extracts the field with the largest value out of the fields entered in this operation.
  ● **AVG**: Calculates the average number of the fields entered in this operation.
● Open a square parentheses “[“ to view a list of the column names on which you can perform actions.

7. To create a conditional expression using the **Case** option:
● Click the **CASE** button. The text field provides an empty expression format.

The expression means that if a certain condition is met, there is a specified outcome. Otherwise, another outcome is expected.

● The condition must be a whole expression in the scope of the WHEN and ELSE parentheses, while the outcome must be a value of the same type.

● Add or delete as many conditions as required. The expressions can also include other sub-expressions within their scope, as long as the limitation of the custom expressions is followed.

● The conditions can include the following operations: =, <>, <, >, >=, <=.

● In addition, the following conditions can be used. These conditions check the following:
  ● IS EMPTY: Checks if a field is empty.
  ● IS NOT EMPTY: Checks if a field is not empty.
  ● CONTAINS: Checks if a field contains a certain value (which must be enclosed in a double quote).
  ● DOES NOT CONTAIN: Checks if a field does not contain a certain value (which must be enclosed in a double quote).
  ● BETWEEN NUMBERS: Checks if a certain numeric field has a value between two numbers.
  ● BETWEEN DATES: Checks if a certain date type field has a value between two dates (which are enclosed in a double quote and written in a YYYY-MM-DD format).
  ● IN: Checks if a certain value exists in a selected column.
  ● AND/OR/NOT: Used to add a relation to the condition.

8. After constructing the expression click the **Validate** icon. A message is displayed indicating if the validation is successful or if there are errors.

9. To save the Custom Expression, click **Save**. To delete the Custom Expression, click **Clear**. After saving it, the new customized column is added to the Column Name list with a Custom Expression indication.
5.2.2. Performing Data Conversion

Data conversion can be used if the user does not want data to be displayed for a specific column (for example, a column with a long list of medication names) but rather wants to display if data from this property exists for the patients in the cohort or does not exist (it will display a True or False value for whether the medication exists in the cohort). In this case, a list of True/False values (or 1/0 values) is displayed in the output file instead of the property values.

The following options can be used for data conversion:

- **Overwriting the values** in an existing column with True/False (or 1/0) values instead of the original values.
- **Creating a completely new column** with converted values (to True/False or 1/0) without overwriting the original column values. In this case, the new column with True/False (or 1/0) values is added to the list of columns.

To convert the data to True/False (or 1/0) values while overriding the original column values:

1. Log in to the MDClone Dashboard. The My Queries screen appears.
2. In a new Query or in a previously created Query, select the Finalize Cohort and Output tab.
3. Select the column for which you want to convert the data and click the Action icon on the right. Then click If Exists Conversion.
   The following Conversion option is displayed on the column row.

   ![Convert to True/False](image)

4. Select the new value you want to display in the column (1/0 or True/False). Note that these values override the previous values in the column.
   For each value in the column, if this value exists for any patient in the cohort, True (or 1) is displayed in the column.
   If this value does not exist for any patient, False (or 0) is displayed.

If a Categorization was applied to this column, the Conversion now overwrites the categorization values. The categorization values will not be displayed in the output file.

The Calculated Column indication is now added to the column row. For example:

![Calculated Column](image)

- Note that when data conversion is performed on a column, the column now is 100% populated in the cohort, even if previously it was populated less than 100%.
  This is because all cells in the column now have either the True/False (or 1/0) value and no cells are empty.
To create a new column with True/False (or 1/0) conversion values without overwriting the original column values:

1. Log in to the MDClone Dashboard. The My Queries screen appears.
2. In a new Query or in a previously created Query, select the Finalize Cohort and Output tab.
3. Click the + Calculated Column button above the Column Name list.
4. From the dropdown menu, select If Exists (True False).
   The Add Calculated Column - If Exist Column screen appears.

5. In the Type a new column name text box, enter a new name for the column.
6. From the If dropdown list, select the column whose values you want to convert to True/False (or 1/0).
   Note that this creates a new column while the original column is not changed or overridden and continues to display the original values.
7. Select True/False or 1/0 values to be displayed in the new column.
   In the following example, a new column was created named MedBeta_Y_N. If a BetaBlocking Drug in the original column exists for any patient in the cohort, its value is True in the new column. If it does not exist for any patient in the cohort, its value is False in the new column.
   In the following example, the new column is named MedBeta_Y_N to differentiate it from the original column that includes the names of all the BetaBlocking Drugs.
8. To save the new column, click **Save**.
The new column is added to the list of columns and the number of columns is adjusted.
The conversion information is added to the **Category/Calculated Column** information in the table.

The cohort is recalculated and the cohort number is displayed at the bottom of the screen.

### 5.2.3. Adding Two-Column Calculation

The **Calculated Column** menu provides the option to create a calculation between two different columns, for example to calculate the delta between the start date of an Admission and the start date of a Procedure (by subtracting the earlier Admission date from the later Procedure date or vice versa). This calculation is then added as a **new column** in the Column list to be generated in the output file.

![Image of calculated column input](image)

**Note:** This calculation can only be applied to columns with a **Numeric** or **Date** value, and the calculation can only be applied between two of the **same data types** (either Numeric or Date).

The calculation between two column values can be one of the following: **addition, subtraction, multiplication, or division.**

**To create a Two-Column Calculation**

1. Log in to the MDClone Dashboard and click **Query Tool > My Queries.** In a new Query or in a previously created Query, select the **Finalize Cohort and Output** tab.

2. Click the **+ Calculated Column** button above the Column Name list. From the dropdown menu, select **2 columns calculation.**
The Add Calculated Column - 2 Columns Calculation screen appears.

3. Enter the following information for the calculation:
   - In the Type new column name text box, type a descriptive name for the new column that you want to create.
   - Select the first column from the dropdown menu.
   - Select the required mathematical Operation from the Operation menu (addition, subtraction, multiplication, or division).
   - Select the second column, from the second dropdown menu.

For example, the following selection performs a calculation on the number of years between the Deceased Date and the PTCA start date.

4. Click Save, to save the calculation.
   The new Calculated Column is created and added to the Column Name List. An indication is added to show that this column represents a 2-column calculation.

Attention! Division by zero or NULL returns Inf as the output.
5.2.4. Adding One-Column Calculation

The user can add additional mathematical calculations to the data in a column by selecting the 1 column calculation option from the Calculated Column dropdown menu. For example, you can calculate the square value of a certain numeric field. This calculation can only be performed on Numeric data.

Each added calculated column is displayed as a new column in the extracted file. The calculations take place before the synthesis process of the data in the file to preserve the accuracy of the calculations.

To add a One-Column Calculation

1. Log in to the MDClone Dashboard and click Query Tool > My Queries. In a new or previously created Query, select the Finalize Cohort and Output tab.

2. Click the + Calculated Column button above the Column Name list.

3. From the dropdown menu, select 1 column calculation.

The One Column Calculation screen appears.

4. Enter the following information for the calculation:
   - In the Type new column name text box, type a descriptive name for the new column you want to create.
● Select a **column** on which you want to perform the calculation from the dropdown menu.

● Select one of the following mathematical **operations** from the **Operation** menu: **Ln** (logarithms) and **POW** (exponentiation).

5. Click **Save**, to save the calculation.

The new Calculated Column is created and added to the Column List.

An indication shows that this column represents a One-column calculation.
6. Managing My Queries

When the user logs in the dashboard, the My Queries page is displayed by default. The following figure shows the My Queries screen, where the user can manage his queries and the queries that were shared with him (under Queries Shared With Me).

To manage My Queries:
1. Log in to the MDClone Dashboard. The My Queries screen is displayed. It displays all the queries the user has performed in the MDClone My Queries platform. The features in My Queries screen are described in My Queries View.
2. To perform additional action on My Queries or while creating a new query, see Performing Additional Actions on Queries.

6.1. Performing Additional Actions on Queries

1. Log in to the MDClone Dashboard and select Query Tool > My Queries. The My Queries screen displays the list of the user’s queries.
2. Select the query on which you want to perform additional actions. For most of the actions, click the Action icon to display the actions that can be performed on the selected query.

Note that the available actions might differ during different stages in creating queries.

- **Save**: The Save option is available while creating a query either by clicking the Save icon or from the Action menu. It is only available when there are unsaved changes.
- **Event Update Notifications**: See Viewing Notifications for Event Updates.
- **Edit**: For details, see Editing a Query.
- **Duplicate**: For details, see Duplicating a Query.
● **Download Files**: To download the query output files, click **Download**. The files are downloaded in a zip file, containing the generated CSV file, an Info PDF file (containing the content of all five steps of the query creation process), and a PDF Comparison report (for Synthetic data).

● **Share**: For details, see **Sharing a Query**.

● **Send**: For details, see **Sending a Query**.

● **Delete**: For details, see **Deleting a Query**.

● **Quality of Care**: For details, see **Quality of Care**.

3. To expand the **Queries Shared with Me** list click the arrow on the left. The Shared Queries list is displayed.

The following information is displayed for **Queries Shared with Me**:

- **Number** of queries in the list.
- **Name**: Name of the query. Click the name to open the Query Tool and either view or edit the query (depending on the granted permissions).
- **Modified**: Date the query was last modified.
- **Shared by**: Name of user who shared the query (see **Sharing a Query**).
- **My Permissions**: Permissions granted for this query, either **View** or **Edit**.

4. Click the **Action** icon to display the available actions that can be performed on **Queries Shared with Me** (for details, see **Sharing a Query**).

- **Edit or View**
- **Duplicate to My Queries**
- **Remove from List**

### 6.1.1. Viewing Notifications for Event Updates

If the Admin updates an Event or Properties settings, the admin can choose to send a notification to the end user that describes the update.

The user can view all the event update notifications by clicking **Updates** in the upper right hand corner of the Query where the events were updated.
The following figure shows an example of query **Update** notifications. It lists the events or properties that were updated.

![Update notifications](image)

### 6.1.2. Editing a Query

**To edit a query:**

1. Log in to the MDClone Dashboard and in the navigation menu, select **Query Tool > My Queries**.

2. To open and edit a Query, click the query’s name or click the **Action** icon on the right of the query and click **Edit**.

   The following information is displayed for the query:
   - Name of the query.
   - Description of the query.
   - Last Saved - Displays the date and time the query was last saved.

3. The query settings and definitions can be edited and changed as described in the five steps in building a query (see **Building a Query**).

### 6.1.3. Sending a Query

You can send your query to other users in your organization to provide them with a private copy of your query. The query is the query created by the sender but only with the permissions of the recipient.

Notice that if you send a query with a Customized Event in its definition, the recipient of the query will have his own private copy of the Customized Event.

**To send a query:**

1. Log in to the MDClone Dashboard and in the navigation menu, select **Query Tool > My Queries**.

2. To send a Query, click the **Action** icon on the right of the query and click **Send**.
The following **Send Query** screen appears.

3. In the **Send to** field, enter or search for the name(s) of the recipient(s) of the query. Notice that due to permission settings, the recipient might not be able to see all the query domains.

4. Click **Send**.

A copy of the query is sent to the defined recipient and the recipient can manage the query in the same way as any query is managed (it can be edited, sent, shared, deleted, etc.).

The following behavior occurs for the recipient:

- When receiving a query, it is displayed in the recipient’s **My Queries** list. The query’s name begins with **FWD**: followed by the original query name. For example: `FWD:PTCA`

- If the recipient does not have permission to see some events/properties in the sent query, an Alert or Warning icon indicates that there is a problem. For example in the **Events List** the following Alert is displayed:
In the **Event Definition** tab, the event is displayed as **Restricted**.

![Event Definition Tab](image)

The following message is displayed:

![Errors Found in Page](image)

The recipient cannot continue to manage the query without resolving the problem, for example, by deleting the Restricted Event.

- The recipient of a query that contains a **Simple Combined Event** receives a copy of the event, that can later be regenerated and redefined.
- The recipient of a query that contains a **Complex Combined Event** receives a locked copy of the event, a fixed cohort, and there is no option to regenerate or redefine the event. The combined event is treated as a Load Event by the platform.

### 6.1.4. Sharing a Query

You can collaborate with other users in your organization by sharing your query. This enables other users to participate as a viewer or editor in your query. However, only the owner of the query can generate an output file.

**To share a query:**

1. Log in to the MDClone Dashboard and in the navigation menu, select **Query Tool > My Queries**.
2. To share a Query, click the **Action** icon on the right of the required query and click **Share**.
   The following **Share Query** screen appears.
3. In the **Share with** field, search for or enter the name(s) of the user(s) with whom you want to share the query.

4. In the **Permission** field, select the permissions you want to give to the selected user. You can select **View** or **Edit** permissions.
   - **View**: The user can view all the events in the query, but cannot edit them.
   - **Edit**: The user can edit all the query settings.

5. Click **Add**.

The user is added to the list of users with whom you are sharing the query.

Change the permissions per user, if required.

6. Click **Share**.

The query is now added to the user’s **Queries Shared With Me** list.

The **Shared** icon is added to the original query in the **Shared** column.
The **Queries Shared with Me** area displays the following information:

- **Number** of queries that were shared with me.
- **Name**: Name of the shared query.
- **Modified**: Date the query was last modified.
- **Shared by**: Name of user who shared the query with me.
- **My Permissions**: The permissions given to me on the query (View or Edit).

7. The following actions can be performed on the Shared Queries by clicking the **Action** icon.

- The displayed actions are either **View** or **Edit**, depending on the user’s permissions.

If the Permissions are defined as **View**, the events in the query cannot be edited. The query can only be viewed.

If the permissions are defined as **Edit**, the user can edit the shared query. All other shared users of the query can see these changes and perform their own changes (in accordance with their permissions).
The actions allowed to shared users are based on the permissions of the query’s owner. Note that depending on organizational policy, administrators can be allowed to edit a shared query regardless of the permissions of the query’s owner.

**Download:** A shared user cannot perform the last step (Step 5) in building a query and cannot download a result file from the query (unless the user is an admin).

**Duplicate to My Queries** option (see Duplicating a Query) enables the user to create a private copy of the query in the My Queries list. This enables the user to manage the query as his own and also download a result file at the end of the process.

**Remove from List:** Select this option to remove the query from the list.

### 6.1.5. Duplicating a Query

**To duplicate a query:**

1. Log in to the MDClone Dashboard and in the navigation menu, select Query Tool > My Queries.

2. To duplicate a Query, click the Action icon on the right of the query and click Duplicate.
   The query is duplicated and added to MyQueries list.
   It is named `<query name> - copy`.

### 6.1.6. Deleting a Query

**To delete a query:**

1. Log in to the MDClone Dashboard and in the navigation menu, select Query Tool > My Queries.

2. To delete a Query, click the Action icon on the right of the query and click Delete.
   After deleting a query it cannot be undone.
   A confirmation message is displayed.

3. Click Yes to delete the query.
   The query is deleted from the My Queries list.
7. Creating Quality of Care

A Quality of Care file is a result file of a query that has multiple occurrences of the Reference Event (meaning, multiple rows of an event) for each patient in the cohort. In a query’s standard result file, there is only one occurrence of the Reference Event (one row) for each patient in the cohort. However, in some cases, the user might want to view all Reference Events for patients (and not only a single occurrence of a Reference Event). For example, an organization might want to compare how many admissions occurred in a specific year compared to all admissions in the previous year. Each occurrence of the Reference Event (each row) in the Quality of Care file includes the relevant Time-Related Event(s) for that occurrence. For example, if there are multiple occurrences of Admission for a patient, each Admission occurrence includes the Time-Related Event(s) for that specific occurrence.

Note that Quality of Care cannot be created for Queries that contain Multiple Occurrence events (see Defining a Multiple Occurrence Event).

Creating a Quality of Care file is only available for the following:
- For queries that already have a result file
- For users with permission to see Original Data (this includes Synthetic Data users who have an Approval Key)

To generate a Quality of Care file:
1. Log in to the MDClone Dashboard and select Query Tool > My Queries.
2. Select the relevant query, click the Action icon, and then click Quality of Care. The Quality of Care screen appears:

   ![Quality of Care Screen](image)

   - The result file will include a row for every occurrence of the Reference Event up to the maximum defined below.
   - Each row includes the relevant Time-Related Events.
   - Maximum occurrences (rows) for a patient [1-99]: 5

   Note:
   - The occurrences listed in the result file are determined by the time of the Reference Event’s occurrence.
   - The creation of the result file will be as scheduled by the admin.
3. Set the maximum number of rows per patient (based on the maximum number set by the admin for Quality of Care). The default value is 5.

4. Click Submit.

The Quality of Care file is submitted and will be generated as scheduled by the admin (this is usually a process that runs overnight; however, the admin can change the setting to Run Now).

The scheduled status is displayed in the following screen:

A summary is provided in a table and shows the following information:

- **Status**: The following statuses might be displayed:
  - **Submitted**: This is the status of the file until it is generated. The file is submitted and will be generated after the generation process runs.
  - **Completed**: The file is successfully generated and can now be downloaded and rescheduled.
  - **Failed**: The file failed to be generated. It is recommended to run it again or contact the admin for help.

- **Max Rows**: The defined maximum number of rows per patient.

- **Submitted on**: The date and time the user clicked Submit.

- **Actions**: The following actions are available (after the file is submitted the only available action is Delete):
  - **Delete**: Cancels the process of generating the file.
  - **Reschedule**: Reschedules the generation process. This option is relevant after a failed process, or when changes were made in a query and the user wants to generate a new Quality of Care file that includes the changes.
  - **Download**: Downloads the Quality of Care file. This option is available after the file is generated (unless the user clicks Reschedule). When the new rescheduled file is generated, the Download option becomes available again.
5. Click **Close**.
When the Quality of Care file is generated a notification message informs the user that the process was completed successfully and the file is ready to be downloaded.

![Quality of Care completed successfully]

6. Download the file by doing one of the following:
   - Click the **Download** link in the notification.
   - Or
   - In the My Queries screen, select the relevant query, click the **Action** icon, and select **Quality of Care**. The **Quality of Care** screen appears. In the **Actions** column click the **Download** icon.

The Quality of Care CSV file is downloaded.

The following is an example of the QOC file name:

```
[8945][QOC][mdclone.admin][qe_2947][20210914_083512275].csv
```

**qe_2947** represents the Query ID.

To view and interpret the Quality of Care file, note the following:

- Multiple rows might belong to the same patient.
- It is recommended to include the patient id in the output file in order to identify the patients with multiple occurrences of the Reference Event.
- Each row describes one Reference Event and its relevant Time-Related Events.
- The result file contains an additional column (column A) that indicates which occurrence it is of the Reference Event (for example, 1, 2, 3, or 4, etc.). The name of the column is **Reference occurrence number**.
8. Managing My Events

The My Events screen enables the user to add and manage a Combined Event or a Loaded Event. Combined Events that were sent to users by other users of the Query Tool platform are also managed here. The events in the list can be used later in the user’s queries and sent to other users.

To access the My Events main screen, click My Events in the navigation pane.

The My Events main screen appears:

![My Events Screen](image)

The following features are available from the My Events main screen:

- Creating a Combined (AND) Event
- Defining a Combined (OR) Event
- Creating a Private Loaded Event
- Viewing My Events List
- Sending an Event

8.1. Combined Event Overview

A user can define an event as two or more combined events in order to more precisely define the cohort when creating a query. When defining the Reference Event, the user generally selects a single Reference Event (for example, the first or second occurrence only of the event). Defining a Combined Event enables the user to define a Reference Event as a combination of two or more events to more specifically define the cohort.
This capability is not available when normally defining the Reference Event (see Step 1: Defining a Reference Event), which is always a single specific event. When defining a Combined Event:

- A Combined (AND) Event can only define a maximum of two events.
- A Combined (OR) Event can define a maximum of four events.

8.2. Creating a Combined (AND) Event

A combined AND event is defined with two events that are related in time. Only patients with both of the events are included. After defining two events (Event I and Event II) the user defines the time relationship between them and generates the event.

The combined event can be defined with events from the organization’s domain (such as Medication and Diagnosis, etc.) as well as with events from the user’s event list (such as other combined events and population events).

To create a Combined (AND) Event:

1. Log in to the MDClone Dashboard, select My Events, and click + Combined Event.

   ![Combined Event Creation Screen]

   The following screen appears where the user can search for the required event.

   2. Type a Name and Description (optional) for the Combined Event.

   3. Do not change the AND Event Type (default).
4. Define **Event I** using the **Search Engine** (see *Defining an Event Using the Search Engine*) or **Manually** (see *Defining a Reference Event Manually*).

   Note that at this point it is still possible to change the event type to **OR** (if required). After defining the second event, this option is no longer available.

   The following example shows that the selected **Event I** was defined as **Admissions**.

5. Add **Event II** by clicking **+ Additional Event**.
6. Define **Event II** using the **Search Engine** (see *Defining an Event Using the Search Engine*) or **Manually** (see *Defining a Reference Event Manually*).

   The following figure shows the definition of **Event I** (Admissions) and **Event II** (Diagnosis).

7. After Event II is defined, define the timing between Event I and Event II. The Timing definition area appears.
For a detailed description about how to set the timing definition between two events, see Setting the Timing Definition of the Time-Related Event.

To define the timing between the events:

a. From the Get dropdown menu, select which occurrence of Event II you require (for a detailed description of the available options, see Defining the Occurrence of a Time-Related Event).

The default option is the first occurrence of Event II.

b. Select the Event II date property from the dropdown menu (for example, the Start date).

c. Select the time relationship between the events from the dropdown menu (whether Event II occurred at, after, before, ever, or between two dates in relation to Event I).

d. Define the date property of Event I (for example, the Start date).
If the user defines a time-relationship that occurs **between two dates** of Event I, an additional dropdown list with a **date property** must be defined.

e. If necessary, use the Slider to set the timing between the events (for a detailed description of configuring the timing of events with the slider, see [Defining a Time-Related Event with Slider](#)).

The following example shows a timing definition between two events, where **Event II Start date** is defined as occurring up to 15 days after **Event I Start date**.

8. **Click Save & Generate.**
The Combined Event is generated and the system returns the user to the My Events screen. While generating the Combined Event the **Generating** icon 🔄 is displayed. If the Combined Event was generated successfully a success notification appears and the Event is added to the list of My Events.

   ![Diagram of timing definition between events](image)

   **Note** that the **Generating** icon continues to be displayed until the user clicks **Refresh**.

9. **To change the icon to the Completed status ✔️, click Refresh.**
It is now possible to create a query using the Combined Event.
For example, when defining the Reference Event (see [Step 1: Defining a Reference Event](#)), the Events dropdown menu now displays **My Events**, including Combined Events.
8.3. Defining a Combined (OR) Event

A Combined (OR) Event can be defined with up to four events (minimum number of events is two). Patients with one or more of the events are included in the cohort. The combined event can be defined with events from the organization’s domain (such as Medication and Diagnosis, etc.) as well as with events from the user’s event list (such as other combined events and population events).

To define a combined (OR) event:
1. Log in to the MDClone Dashboard and select My Events.
2. In the My Events screen, click + Combined Event.
3. Type a Name for the event.
4. Type a Description of the event (optional) in the next field.
5. Change the selection of the Event Type to OR.
6. Define Event I using the Search Engine (see Defining an Event Using the Search Engine) or Manually (see Defining an Event Manually). Note that at this point it is still possible to change the event type to AND (if required). After defining the second event, this option is no longer available.
7. Add Event II by clicking + Additional Event.
8. Define Event II using the Search Engine or Manually.
9. Add and define Event III and Event IV, if required. A minimum of two events is required to generate a Combined OR Event.
10. Click Save & Generate.
The Combined Event is generated and the system returns the user to the My Events screen. While generating the Combined Event the **Generating** icon 🔄 is displayed. If the Combined Event was generated successfully a success notification appears and the Event is added to the list of My Events.

Note that the **Generating** icon continues to be displayed until the user clicks **Refresh**.

10. To change the icon to the **Completed** status ✔️, click **Refresh**. It is now possible to create a query using the Combined Event. For example, when defining the Reference Event (see [Step 1: Defining a Reference Event](#)), the Events dropdown menu now displays **My Events**, including Combined Events.

Note: When defining a Reference Event based on a Combined OR Event, there might be a few Time-Related Events: for the customized event itself and for every event defining the customized event (up to four events).

### 8.4. Creating a Private Loaded Event

The user can load and configure an Event in an environment by uploading it from a CSV file. This Event can later be used when creating a query. For example, a user might require only specific Events and data to be part of a query. These events are **private events** and **are not available to other users** when creating a query.

When the user uploads a Loaded Event, the user uploads a CSV or XLS file, maps its columns, and checks the match between patients in the file and patients in the organization’s database.

If the Event loading is successful, only the patients that exist in the organization’s data lake (demographic table) are uploaded to the system. The file must contain the Patient IDs and the Event Date.

The file must comply with the following requirements. These requirements are also listed in a tooltip when uploading the file.

- File type must be CSV or Excel.
- The column headers must be unique (non-identical) and cannot contain an empty value.
- Contact your admin for more information about organizational privacy policy:
  - Patient IDs should comply with the organizational privacy policy.
  - Dates should be shifted according to organizational privacy policy before upload.
- For files with no date column, choose either the **Loading Date** or the **Patient Birth Date** as the event date.
The following Date formats are supported:
- MM/DD/YY HH:MM:SS
- DD/MM/YYYY HH:MM:SS
- YYYY/MM/DD HH:MM:SS

**Note**! Millisecond information is not supported.

**To view and add Private Loaded Events:**

1. Log in to the MDClone Dashboard and select **Query Tool > My Events**.
2. In the **My Events** screen, click the **Loaded** tab.
   
The Loaded Events list is displayed.

The Loaded Events list displays the following information for each uploaded file:
- **Name** of the uploaded file
- **Created**: Creation date
- **Uploaded**: Date the file was uploaded
- **Columns**: Number of columns in the Loaded Event
- **Rows**: Number of rows in the Loaded Event

Click the **Action** icon to view a dropdown list with additional available actions.

3. To add a loaded event, click **+Loaded Event**.
   
The **Upload File** screen appears.
4. Choose a file to upload and click **Upload File**.

If the upload is successful, a Success message is displayed.

![Upload File](image)

A summary of the file’s row and column count is displayed under the **Choose file** field.

File uploaded successfully with 12 columns and 6,631 rows

5. Select the appropriate column headers in the uploaded file to match the following fields:
   - **Patient Key**: Select a column in the uploaded file that matches the IDs of the uploaded population.
   - **Event Date**: Select a column in the uploaded file that matches the date of the event. The Event Date must be one of the two columns that are automatically added to the file: **Loading Date** or **Birth Date**.

6. To check whether the data in the selected columns matches your organization’s data, click **Check Match**.

A **Loaded Event Match** notification is displayed.

![Loaded Event Match](image)
It displays the following percentages of matching for the two fields:

- **Patient Match**: Percentage of patients in the uploaded file that match the patients in the organization’s demographic data. For example, your file might include additional patients who are not included in your environment. In this case, the matching is not 100%.

- **Event Date Match**: Percentage of the data in the Event Date column that matches the data in your organization.

- It also displays the number of rows from the file that will be uploaded.

In the example, the Match notification example shows only 95.01% matching for the two fields. As a result, only 476 (out of 501 rows) rows in the file will be uploaded.

7. If the matching results were not high it is recommended to try to map other columns from the file. To do so, click **Map Again** and select different columns to map to the Patient Key and Event Date.

8. To view the rows with unmatched patients or dates, click the **Click here** link.

A CSV file is downloaded with the list of mismatched patients and dates.

9. If the matching is satisfactory, click **Continue**.

The Matching now displays a successful **Matched** indication.

10. Enter the information in the following fields:

- In the **Event Alias** field, type a name to identify your new Event.
- In the **Description** field, type a description for your event. It is displayed for end users in a tooltip in the Query Tool.

11. Click **Next**.
The **Define Columns Type** screen appears and displays the Column Names, the configured Keys, the Data Types, and Data Samples in the uploaded file.

The **Data Types** can be changed to String, if required.

12. To upload the Event file, click **Upload Event**.

The **Load Summary** screen appears and displays a summary of the event loading.

The following data is displayed:

- The name of the event and the name of the table created in the database.
- The statistics of the event loading:
  - The percent of the patients from the file that match the organization’s demographic data.
  - The number of rows from the file that are uploaded.

13. To save the uploaded event, click **Save and Go to My Events**.

The **My Events** screen appears with the uploaded Event added to the list of Events. The Event can now be used when creating a Query.

14. To update a Loaded Event, follow the steps in **Updating a Loaded Event**.
### 8.5. Updating a Private Loaded Event

The following steps describe how to Update an existing Private Loaded Event rather than creating and uploading an entire new file.

**Note that updating a Private Loaded Event is only available for events loaded from version 6.1 but not for events loaded prior to version 6.1.**

To update a Loaded Event:

1. Log in to the MDClone Dashboard and select **Query Tool > My Events**.

2. In the **My Events** screen, click the **Loaded** tab.

   The Loaded Events list is displayed.

3. To perform additional actions on an uploaded file, click the **Action** icon to the right of the file. The following actions on Loaded Events are available:

   - **Update**: Updating a Loaded event is described in the following steps
   - **Send**
   - **Delete**

4. To update a loaded event, select the required Loaded Event. In the Action dropdown list, click **Update**.

   The **Update Loaded Event** screen appears.
5. The following guidelines and limitations are displayed in the right hand area of the screen when selecting to **update** a loaded event. Make sure to upload a file that **conforms to these guidelines**!

- Columns can be added but cannot be deleted.
- The Column Names must not be changed.
- Data can be added or changed.

6. To update the Event by editing the Event **Alias** name or the Event **Description**, enter the changes in the relevant fields and click **Save**.
7. To update the Event by overriding it with a different file (for example, a file with additional columns or with additional rows for added patients), select the Choose a file for upload option and browse to the CSV or Excel file that contains the updates you require.

The new file must conform to the provided guidelines for uploading an updated file. This file will completely override the existing Loaded Event file and replace it.

8. Select the required file and click Open.

If the file is successfully uploaded, a Success indication is displayed.

9. The Patient Key and Event Date columns are disabled and cannot be changed. These columns must match the previous Load Event file.

Click Check Match to verify that these columns match the columns in the overridden Load Event file.

If matching is successful a success message is displayed.
10. Click **Continue** to return to the Update Loaded Event screen and follow the steps as described in *Creating a Private Loaded Event*. Note that in the **Define Columns Type** tab, the column type cannot be changed in previously existing columns. Adding a Loaded Event
It is possible to change the **Data Type** only in a new column. A newly added column will display a **New** indication.

11. To download the event as a file that existed before it was updated, click the **Download File** icon in the upper right hand corner of the screen. The event is downloaded as a CSV file. If required, you can edit and make changes in the file and upload it again.

### 11.1. Viewing My Events List

The My Events list displays the following information about each Combined Event that the user defines, generates and uploads.

**To view My Events:**

1. Log in to the MDClone Dashboard and select **Query Tool > My Events**. The following information is displayed:
   - **Name**: Name of the event.
   - **Modified**: Date when the event was last modified.
   - **Generated**: Date when the event was last generated.
   - **Event Type icon**: Displays one of the following Event Type icons:
○  o: Combined event type OR
○  : Combined event type AND
○  : Uploaded Loaded Event

●  **Status:** displays one of the following Event Statues:

○  ✔ **Completed:** The event was successfully generated/uploaded.

○  □ **Failed/Error:** The following are possible Fail/Error reasons that are listed in the Event List:
  ■  Event is empty: There are no patients in the organization’s database that match the combined event definition.
  ■  Combined Event generation failed.
  ■  Combined Event exceeded size limitation: The size limit is defined by the admin. The final size of the combined event exceeds this size. Try narrowing its definition by adding filters to the event or contact the admin.

○  ! **Alert:** Other combined event/s from the Event List are used to define this event. The inner event is more updated than this event. It is recommended to regenerate this event.

○  ○ **Generating:** The event is currently being generated.

○  ○ **On hold:** This event was sent by another user or is used in a query that was also sent by another user. This event was not generated yet.

2.  **Action icon:** to perform additional actions on the Events, click the Action icon (the available actions depend on whether the event is a Combined Event or a Population Event):

○  **Edit:** Enables the user to edit the Combined Event.

○  **Duplicate:** Creates a copy of the event in the list.

○  **Send:** Enables the user to send a copy of the event to another user in the organization (see Sending an Event). Note that the user might not be able to see or use the event due to permission differences.

○  **Delete:** Clicking Delete, displays a confirmation message. To delete the file, click Yes. Notice that if a combined event is being used in a query or in another event, it cannot be deleted unless the user deletes its references.

3.  In the header of the screen an **Action icon** provides the following additional options. The availability of the actions depend on whether the event is new or an existing event:
○ **Save & Generate:** This option is available only for an existing event; otherwise it is disabled. It enables the user to save the recent changes and regenerate the event.

### 11.2. Sending an Event

The user can send a copy of the event to a selected user from the organization. Notice that the recipient of the event might not have permissions to use the event’s components. In this case, parts of the event are labeled as Restricted.

**To send an event:**

1. In the My Events screen, select an event, click the **Action** icon and choose **Send**. The **Send Event** window appears.

   ![Send Event Window]

   - The selected event will be copied internally to the recipient event list. Please note that the recipient may not have the permissions to use the event components.

2. In the **Send to** field, search for or select the name of the recipient from the list.
3. Click **Send**.

**Receiving an Event (Recipient Side):**

The following behavior occurs when a user receives a query with a private event from another user:

- When receiving an event, the Event appears in My Events list with **On Hold** status. The received event is not generated until the recipient uses the event in a query or enters the event and clicks **Regenerate**.
- When receiving a Combined Event that is defined with other combined events, or a Population Event, the recipient cannot edit or regenerate it. The event appears as **locked**.
12. My Permissions

The My Permissions screen displays a summary of the permissions given to the user by the admin.

The following figure shows an example of the My Permissions screen.

![My Permissions Screen]

To view My Permissions:

1. Log in to the MDClone Dashboard and select My Permissions.

   The My Permissions screen displays the following information:

   - **Allowed Permissions Header:** The following information is displayed in the upper area of the screen.

     ![Allowed Permissions]

     - **Allowed Cohort Size:** The maximum number of patients allowed in a cohort.
     - **Allowed Date Range:** Only events that occurred during this time frame can be queried. The time frame can be configured as between two specific dates, or one or both dates can be set as Unlimited.
For example:

<table>
<thead>
<tr>
<th>Allowed Date Range:</th>
<th>Unlimited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allowed Date Range:</td>
<td>Feb 2, 2000 to Feb 2, 2018</td>
</tr>
<tr>
<td>Allowed Date Range:</td>
<td>Unlimited to Feb 2, 2018</td>
</tr>
</tbody>
</table>

- **Allowed Events**: Lists the Events that the user is allowed to query (as defined by the admin). Only these events are available when defining an event in a query. Events that are not listed but were received in a sent event or a query (see [Sending a Query](#)) are labeled as Restricted Events. The following information is displayed for each allowed event:
  - **Group**: Name and icon of the event’s Display Group (if a Display Group was configured).
  - **Event Name**: Name of the event.
  - **Event Details**: Move the cursor over a row to display the Event Details. The Details pane shows the following information:
    - **Description**: A description of the event as provided by the admin.
    - **More details link** ([More details](#)) (if available): Enables the user to access a PDF file with additional information.
    - **The available data is since**: The earliest year from which data in the event is available.
  - **Approval Keys**: Displays the user’s Approval Keys and the number of Approval Keys. A user can have one or more Permission Sets defined by the admin. Each Permission Set can be configured with an Approval Key.
    - **Key Name**: Name of the Approval Key given by the admin.
    - **Permission Set Name**: Name of the Permission Set where the Approval Key is defined.
    - **Expiration Date**: The Approval Key’s expiration date as defined by the admin.
    - **Status**: Displays the status of the Approval Key:
      - ![Active](#): An **Active** Approval Key allows the user to query original data.
      - ![Expired](#): An **Expired** approval key allows the user to query only synthetic data.
13. Importing Files

The **Imported Files** view enables the user to upload and use external CSV files in the Query Tool platform. These files can then be synthesized using MDClone’s synthetic engine. When an imported file is uploaded to the system, the imported file is **Not Shared** (private) by default, meaning that other users in the organization cannot access the file. During editing, the user can change the setting so the file is **Shared**; however, after the file is mapped, the Share setting cannot be changed.

Note that an admin has access and ability to delete all the imported files that have been uploaded by other users, even files that are labeled as Not Shared.

13.1. Viewing Imported Files

To view the list of Imported Files in your organization, click **Imported Files** in the navigation pane.

To view the list of imported files:

1. Log in to the MDClone platform as an administrator and in the navigation pane, click **Imported Files**.
   The **Imported Files** main window appears.

2. Filter the list of imported files by using one or more of the following filters:
   - **Time Filter**: Filters the file list to show only files modified during the selected time range.
     Select to view All files, or only files modified in the Last Month or Last Week. Selecting All, cancels all the filters.
   - **Environment Filter**: Filters the list to show files either in all or in selected environments.
   - **Permission Template Filter**: Filters the list to show files with only selected Permission Templates.
   - **Search**: Filters the list to show files only containing the string entered in the Search box.
The Imported Files list displays the following information about each file that was imported to the platform during the selected time frame (or according to search results), and includes the following information:

- **Name**: Name of the imported CSV file.
- **Modified**: Date when the file was last modified.
- **Environment**: Name of the environment to which the file was imported. This can limit users to only the specified template.
- **Permission Template**: Name of the Permission Template, if one was selected during import. This can limit users to only the specified template.
- **User Name**: Name of the user who imported the file.
- **Shared**: By default, imported files are NOT shared, but the user can select the Share option when importing the file (see Importing a File). If the Share option was selected during import, the Share icon is displayed and indicates that the file is a shared file.
- **Columns**: Number of columns in the imported CSV file.
- **Rows**: Number of rows in the imported CSV file.
- ****: Indicates that file mapping is required if file mapping was not yet done.

To map the file, click the Action icon and select Edit File Settings. The File Mapping screen appears (see File Mapping).

3. To perform additional actions on a file, select the file and click the Action icon. The available actions depend on whether the file is final and if it is shared or not shared.
   - **Edit File Settings**: If enabled, the user can edit the file settings (file mapping).
   - **View Output Summary**: Displays the Output Summary. Enables the administrator to use a synthetic data view mode when working on an imported file (see Step 5: Finalizing the Cohort and Output).
   - **Delete**: Clicking Delete, displays a confirmation message. To delete the file, click OK.

### 13.2. Importing a File

To upload and import a CSV file to the MDClone platform:

1. Log in to the MDClone platform as an administrator and in the navigation pane, click Imported Files. The Imported Files main window appears.

2. In the Imported Files window, click + New File . The New Imported File window appears.

3. In the Choose File field, click Choose File to select the file to upload. The selected file is then shown in the text box.
4. In the **Name** field, type a name for the file.
5. In the **Environment** field, from the dropdown menu, select the environment to which you want to upload the file.

6. **Sharing**: By default, the file is Not Shared. A file that is private (not shared) is only accessible to the file owner and also to his admin. In addition, a file that is Not Shared can only be deleted by the owner or by his admin. To enable the file to be Shared (to be public), change the **Share** option to **Yes**. The file is changed to be a Shared File and can be used by all users in the environment. The following is displayed:

7. If required, select a **Permission Template** (optional) from the dropdown list. Each template includes the group of events that the user can query. Selecting a Permission Template limits the uploaded file only to users with these permissions.

8. Click **Upload**. The file is uploaded and is displayed in the **Imported File** List. If the upload was successful, a Success message is displayed.
9. To complete the upload, click the **Map File** link, to open the **File Mapping** screen. Perform file mapping as described in the next section (**File Mapping**).

### 13.3. File Mapping

After uploading a file, the file must be mapped in the **File Mapping** screen, in order for the uploaded data to be used within the environment.

**To perform the File Mapping:**

1. Do one of the following:
   - Click the **Map File** link in the Success Message displayed after uploading a file or click the **Notification** icon to display the **Map File** link.
   - In the Imported File list, select the file, click the **Action** icon for the uploaded file, and click **Edit File Settings**. The **File Mapping** screen appears.

   ![File Mapping Screen]

2. The following information and options are available from the **File Mapping** screen:
   - **File Name** is displayed at the top of the screen. For example:

     ![File Mapping Screen](Image)

     Change or edit the file name, as required.
● Additional information about the file is displayed at the top of the screen. The following figure shows information that was entered during the file uploading process, as well as the Date Modified, Number of file Rows and Columns.

Environment: Demo Project | Not Shared | Modified: Jun 20, 2021 | Columns: 8 Rows: 17554

3. To change the **File Sharing** setting click the **Shared (or Not Shared)** heading to open the **File Sharing** dialog box.

   ![File Sharing Dialog Box](image)

   Change the setting to **Share (Yes)** or **Not Share (No)** the file, if required. Then click **Apply**.

4. If the file is **Shared**, select a **Permissions Template** from the dropdown list (optional).

5. The File Mapping table displays the following information for each column in the file:
   - **Checkbox** for each column: Select the columns for which you want to save the mapping. Only the selected columns are saved.
   - **Column Name**: Name of each column in the file.
   - **Data Type**: The data type in the column (String, Numeric, or Date).
     If required, change the **Data Type** from numeric or date to string by selecting the string icon.
   - **Categorization**: The categorization of some columns can be selected from the dropdown list. If selected, only data in the selected category is available.
   - **Percentage Match**: Shows the percentage of data that matches the selected categorization. The following example shows 0% matching:

   ![Percentage Match Table](image)

   - **Data Sample**: Shows an example of the data in the column.

6. After completing the mapping, click **Save**.
   A Progress message appears showing the mapping progress.
When File Mapping is completed, the file is imported to the system.
14. My Output History

My Output History report lists the end user’s output activities (creating Queries or importing files). Each Query can be downloaded as a ZIP folder, including all the files which the user has extracted (customized, original file, info file, synthetic file, multiple occurrence event file, and comparison report), files which were created by the application and a CSV file with patient data.

To access the report, click **My Output History** in the left hand navigation pane. **My Output History** report is displayed.

![My Output History Report](image)

Filter the report by selecting one of the following options:

- **All**: Displays all activities of all dates
- **Last 30 Days**
- **Last 7 Days**
- **Specific**: The user can select a specific date range by using the calendar icon.
- **Search**: The Search function can be used to filter the list for specific data.

**My Output History Report** displays the following information:

- **Date**: Date the output was created (either by a query or file import). All activity is displayed up to now.
- **Query / File Name**: Name of the Query or File Name (for imported file)

Click the link to open the Query in the Query Tool.
● **Environment**: Name of the environment where the output was created. The environment can be filtered, using the **Filter** icon.

● **Data**: Output is either Original Data or Synthetic Data

● **Columns**: Number of columns

● **Rows**: Number of rows

● **Status**: Status of the output file. The status can be Success ✅ or Failure ⚠️ to generate the output file.

● **Download link**: The link can be clicked to download the generated file. It is downloaded in a zip file. The folder contains a CSV file and also an Info PDF file that summarizes the query definition. It provides a summary of all five steps of the query creation process.

Click the **Download Report** arrow to download the report as a CSV file.
## 15. Appendix A - Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Censored Data</td>
<td>In cases where there are too many unique values that may cause patients to be recognized, in order to comply with privacy limitations, the system now censors some value attributes instead of filtering them out. However, even when filtering some values, it is guaranteed that over 98% of the patients will be used for the synthesis. In this way, the user will get a larger cohort with additional essential info.</td>
</tr>
<tr>
<td>Comparison file</td>
<td>The synthetic data processor generates a comparison file that provides a validation for the synthetic process.</td>
</tr>
<tr>
<td>Event</td>
<td>A clinical fact with a timestamp (for example, medication, procedure, diagnosis) a patient has in the database.</td>
</tr>
<tr>
<td>Categorization</td>
<td>A category is a system of grouping based on a known medical classification tool (for example, commonly used categories in healthcare organizations are ICD 9/10, RxNorm, ATC, and SNOMED). Some of the events are defined by properties that are linked to categorizations.</td>
</tr>
<tr>
<td>Imported Files</td>
<td>It is possible to upload data files from external sources to be synthesized using MDClone's synthetic engine.</td>
</tr>
<tr>
<td>Property</td>
<td>For each event, a property is an attribute used to define the event (for example, admission date, lab results, etc.).</td>
</tr>
<tr>
<td>Query</td>
<td>A query is the representation of a question within the medical organization. It is defined with a patient cohort and its output columns.</td>
</tr>
<tr>
<td>Synthetic data</td>
<td>The MDClone engine can examine medical data, extract their statistical characteristics and dependencies, and extrapolate from these to generate a new data set containing synthetic medical data for fictitious patients. This enables the organization to overcome the risk of public disclosure of medical data and thus to have a zero-risk of patient privacy.</td>
</tr>
</tbody>
</table>
# 16. Appendix B - User Permissions

MDClone’s Dashboard displays data differently for each of the following types of users in the system. The following table shows some of the main differences between the different users.

<table>
<thead>
<tr>
<th>Users Allowed to View Original Data OR Users Allowed to View Synthetic Data with Approval Key</th>
<th>Users Allowed to Only View Synthetic Data (without Approval Key)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can download output file regardless of cohort size</td>
<td>Can download output file only with a minimum cohort size of 200 (200 rows)</td>
</tr>
<tr>
<td>Can create Quality of Care</td>
<td>Cannot create Quality of Care</td>
</tr>
<tr>
<td>Displays the exact patient cohort size.</td>
<td>Displays an approximate patient cohort size. If the size is less than 1000, displays the size range only: Cohort: 685-691. If the size is greater than 1000, displays an approximate size: Cohort: ~36,515</td>
</tr>
<tr>
<td><strong>Time-Related Events screen (left pane)</strong> displays the actual percentage that each Event is populated in the database.</td>
<td><strong>Time-Related Events screen (left pane)</strong> displays the percentage range that each Event is populated in the database.</td>
</tr>
<tr>
<td>Can create a Multiple Occurrence event.</td>
<td>Cannot create a Multiple Occurrence event.</td>
</tr>
</tbody>
</table>

The **Finalize Cohort and Output** screen (Step 5) is displayed differently for different users. For details, see: [Step 5: Finalizing the Cohort and Output](#)
The **Finalize Cohort and Output** screen (Step 5) displays the exact percentage that a column (event property) is populated in the database. This is relevant for an Original Data user who selected the Synthetic Mode.

The **Finalize Cohort and Output** screen (Step 5) displays the percentage range that a column (event property) is populated in the database.

![Percentage Range](image)

<table>
<thead>
<tr>
<th>% Populated</th>
<th>1%</th>
<th>10%</th>
<th>19%</th>
<th>14%</th>
</tr>
</thead>
<tbody>
<tr>
<td>95-100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95-100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95-100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The **Finalize Cohort and Output** screen (Step 5) displays the exact percentage of censored rows in each column.

The **Finalize Cohort and Output** screen (Step 5) displays the approximate percentage of censored rows in each column.

- 97%
17. Appendix C - Decryption Solution

17.1. Decryption Solution Overview

This chapter describes the decryption process for projects that load deidentified data, which includes decryption of patient ids, shifted dates, and retrieval of date values.

An Encryption table is used to contain encrypted keys for ids and date shift. This table is linked to a specific permission template so that only permitted users can access it (meaning that the admin must assign the user to this specific template).

17.2. Getting Decrypted Data

For a user to get decrypted data:

1. Add the encryption_table as an event to the query results.
2. In Step 5 of building a query, select the Open With option.

3. Select the decryption app (see the following figure).

![Decryption App](image)

**Note:** This is not supported for Quality Of Care queries.

Note that the following message is displayed to a user that does not have permission to use the Deidentified Data application.

```
You do not have permission to use this application. Please contact your administrator.
```
In the decryption process, the date columns are replaced with the original dates with a _decrypted suffix ([date]_decrypted). The decrypted patient id column that contains the original id is named Decrypted Patient Id and is added at the end of the file.

Note the following when querying the encryption_table:

- The encryption_table must be set as Get first ever (since the event date has no meaning).
- In the Time-Related Event step do not change the alias of the output columns. The column names must not be changed in the query.
- The decryption process does not work in Synthetic Mode.
- If the encryption column (id or date shift) has missing values the process will fail. This can occur if the event was queried incorrectly (for example, Get first ever was not selected) or if there is a problem in the table (for example, a Patient ID that appears in the cohort and does not appear in the encryption_table).
- If the encryption_table was loaded with a new seed after calculating the session, the process will fail until recalculating the encryption_table.
- The Decryption output file is downloaded without the encryption columns (encrypted_id and encrypted_date_value) since they are confusing to the user and because they might disclose the encryption process to a user.

17.3. Date Value Retrieval

Users with permission to the Encryption table template also have access to the Date Value Retrieval option. This app enables the permitted user to get the date_value for a specific date in order to use it during the Query.

To get a relevant “date_value”, select Date Value Retrieval from the navigation menu, choose the relevant date, and click Get Date Value.

The relevant date value will be shown and automatically copied to the clipboard.

Note: The Date Value Retrieval is displayed in the menu only for permitted users, and only if it was configured to be shown (by default the configuration is false and it does not appear in the menu).