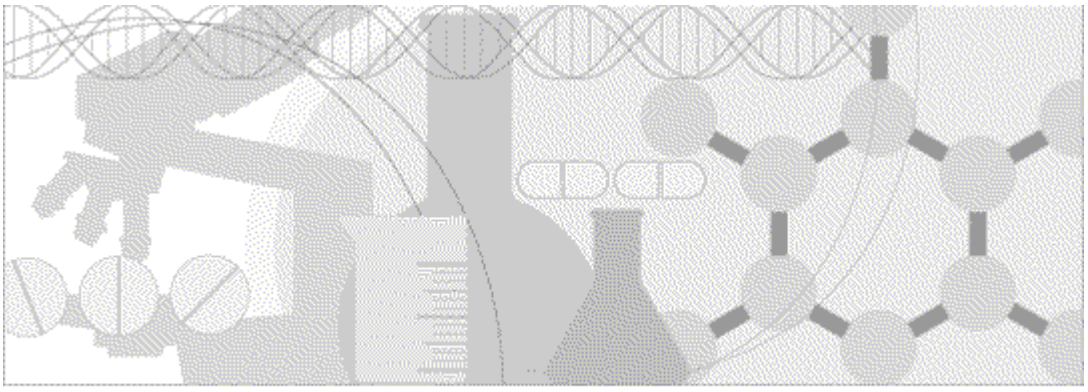
User Guide

Oracle® Health Sciences Empirica Healthcare Analysis Release 1.0.1



ORACLE

Part number: E59664-01

Table of Contents

About this guide

Overview of this guide

The *User Guide* describes how to use the Empirica Healthcare Analysis application to perform epidemiologic and statistical analyses of commercially available healthcare and administrative claims data.

Audience

This guide is for everyone who uses the Oracle Health Sciences Empirica Healthcare Analysis application, including safety specialists, pharmacoepidemiologists, and researchers within pharmaceutical or regulatory organizations, who perform epidemiologic and statistical analyses of commercially available healthcare and administrative claims data.

Documentation

The product documentation is available from the following locations:

- **Oracle Software Delivery Cloud** (https://edelivery.oracle.com)—The complete documentation set.
- **My Oracle Support** (https://support.oracle.com)—*Release Notes* and *Known Issues*.
- **Oracle Technology Network** (http://www.oracle.com/technetwork/documentation)—The most current documentation set, excluding the *Release Notes* and *Known Issues*.

All documents may not be updated for every Empirica Healthcare Analysis release. Therefore, the version numbers for the documents in a release may differ.

Document	Description
Release Notes	The *Release Notes* document provides high-level descriptions of the main features in this release of the Empirica Healthcare Analysis application, as well as system requirements.
Known Issues	The *Known Issues* document provides detailed information about the known issues in this release, along with workarounds, if available.
User Guide	The *User Guide* describes how to use the Empirica Healthcare Analysis application to perform epidemiologic and statistical analyses of commercially available healthcare and administrative claims data.
Installation Guide	The *Installation Guide* describes how to install this release of the Empirica Healthcare Analysis software.
Secure Configuration Guide	The *Secure Configuration Guide* provides guidance and recommendations on securely installing, configuring, and managing the Empirica Healthcare Analysis software and its system components.

Documentation accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc.

If you need assistance

If you need assistance

Oracle customers have access to support through My Oracle Support. For information, visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info, or if you are hearing impaired, visit http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs.

Finding Empirica Healthcare Analysis information and patches on My Oracle Support

Finding Empirica Healthcare Analysis information and patches on My Oracle Support

The latest information about the Empirica Healthcare Analysis application is on the Oracle Support self-service website, My Oracle Support. Before you install and use the Empirica Healthcare Analysis application, check My Oracle Support for the latest information, including Release Notes and Known Issues, alerts, white papers, bulletins, and patches.

Creating a My Oracle Support account

You must register at My Oracle Support to obtain a user name and password before you can enter the site.

1. Open a browser to https://support.oracle.com.
2. Click the **Register** link.
3. Follow the instructions on the registration page.

Finding information and articles

1. Sign in to My Oracle Support at https://support.oracle.com.
2. If you know the ID number of the article you need, enter the number in the text box at the top right of any page, and then click the magnifying glass icon or press **Enter**.
3. To search the knowledge base, click the **Knowledge** tab, and then use the options on the page to search by:
   - Product name or family.
   - Keywords or exact terms.

Finding patches

You can search for patches by patch ID or number, product, or family.

1. Sign in to My Oracle Support at https://support.oracle.com.
2. Click the **Patches & Updates** tab.
3. Enter your search criteria and click **Search**.
4. Click the patch ID number.

   The system displays details about the patch. You can view the Read Me file before downloading the patch.
5. Click Download, and then follow the instructions on the screen to download, save, and install the patch files.

Finding Oracle documentation

Finding Oracle documentation

The Oracle website contains links to Oracle user and reference documentation. You can view or download a single document or an entire product library.

Finding Oracle Health Sciences documentation

For Oracle Health Sciences applications, go to the Oracle Health Sciences Documentation page at http://www.oracle.com/technetwork/documentation/hsgbu-clinical-407519.html.

Note: Always check the Oracle Health Sciences Documentation page to ensure you have the most up-to-date documentation.

Finding other Oracle documentation

1 Do one of the following:

- Go to http://www.oracle.com/technology/documentation/index.html.
- Go to http://www.oracle.com, point to the Support tab, and then click Product Documentation.

2 Scroll to the product you need, and click the link.

Legal notices

This documentation may include references to materials, offerings, or products that were previously offered by Phase Forward Inc. Certain materials, offerings, services, or products may no longer be offered or provided. Oracle and its affiliates cannot be held responsible for any such references should they appear in the text provided.

Getting started

About the Empirica Healthcare Analysis software

The Empirica Healthcare Analysis software is a data analysis environment for exploring multiple sources of the following types of data:

- Population-based temporal, clinical data, such as electronic healthcare records.
- Administrative data, such as insurance claims.

The application informs and supports pharmacovigilance, pharmacoepidemiological, and risk management activities at the following organizations:

- Life sciences organizations, such as pharmaceutical and drug safety regulatory organizations.
- Healthcare organizations, such as payer and provider organizations.

Tasks in a life sciences organization

- Verify safety signals, such as the comparison of adverse events with a drug to those observed with other drugs in the same class or with established therapies for the same indication.
- Identify cohorts of patients based on sophisticated temporal patterns of data, where temporal parameters include specified lead-in periods and risk windows.
- Review prevalence of drug exposures, diagnoses, or procedures.
- Review incidence rates of selected outcomes during specified risk windows.
- Assess risk distribution based on demographic characteristics or time categories.
- Identify most frequently co-occurring terms (such as concomitant drugs, comorbidities, or drug history) or search for the presence or absence of specific co-occurring terms.
- Perform analyses to support retrospective cohort studies, where the occurrence of an outcome is compared for two groups, each of which has a different drug exposure.
- Use demographic variables or history windows as covariates in regression models.
- Use various statistical models to assess risk, including Poisson regression, logistic regression, and Cox proportional hazards model.
- Use high-dimensional propensity scores in logistic regression modeling in order to adjust for differences in the treatment and comparator populations.
- Support strategic decisions regarding the risk management plan for a product.

Tasks in a healthcare setting

- Conduct comparative effectiveness research.
- Explore statistics related to health care services and utilization measurements.
- Detect medical errors, such as co-prescription of medications with known adverse interactions.

- Improve quality of care; for example, by comparing treatment patterns and outcomes before and after a change in policy.
- Identify drug prescription and usage patterns among varying patient populations.
- Identify disease patterns and procedure utilizations.
- Assess whether the difference between groups (a group exposed to a treatment compared to a group exposed to a comparator) is the result of the treatment assignment or of a circumstance unrelated to treatment.

Tasks for all users

- Access large databases of longitudinal data, where you can specify temporal patterns and view both summary and detailed data on the selected patients.
- Explore the characteristics of the patient population before performing analysis.
- Create saved groups of patients to use in further analysis.
- Define and run reports as needed to generate summary or detailed information and graphs from source data.
- Use graphical techniques within the product.
- Download data to study the results of analysis.

Types of data

The Empirica Healthcare Analysis application can support the analysis of the following types of data.

Term in the application	Description
Drug	Drug exposures based on prescriptions and days supplied.
Diagnosis	Medical events, such as diagnoses, symptoms, or conditions.
Procedure	Medical procedures.

You can use demographic data, such as gender or age group, to break down analysis results.

Demographic data also includes dates that represent the start and end of patient enrollment periods, where an enrollment period is the period of time during which the patient could have contributed data to the healthcare database.

When the application is set up, an administrator determines the values in the source data that are used for each type of data. For example:

- Drugs can be set up to use the date on which a prescription was written or the date on which it was filled.
- Enrollment periods can be set up as the longest period of eligibility in an insurance plan, or as the start of the first eligibility period and the end of the last eligibility period.

Configurations

All data is stored in source tables. To control access to the source data, an administrator creates one or more configurations. Each configuration represents source data in a specific way. For example, suppose your source data includes the dates on which a prescription was written and filled. One configuration can be set up with the date written as the start of prescriptions and another configuration can be set up with the date filled as the start of prescriptions.

A configuration determines the following information:

- The drug, diagnosis, and procedure variables that are available when you perform tasks, such as creating an analysis run or a patient query.
- The columns in the source data that are used to determine patient enrollment start and stop dates.
- The variables that are available for defining breakdowns or covariates in analysis runs.
- The hierarchies of terms that are associated with drugs or diagnoses.
- The patient information that is available in drilldown displays.
- The variables that are available in the Browse Data feature.

Logging in

Do not run more than one session of the Empirica Healthcare Analysis application at the same time on your computer, either in a multi-tabbed internet browser or multiple instances of an internet browser.

Prerequisites

Configure the Microsoft Internet Explorer browser for printing color and downloading.

To log in:

1. In the Microsoft Internet Explorer browser, navigate to the URL provided by your site administrator.

   Note: To easily return to the page in the future, bookmark the login page.

2. If this is your first time logging in, in the **Username** and **Password** fields, type the values given to you by your site administrator. Passwords are case-sensitive; user names are not.

3. Click **Log In**, or press **Enter**.

Note: Oracle recommends changing your password after your first login. After logging in, consider setting your user preferences, including the page that appears after logging in.

Related topics

Navigating the application

Exiting the application

Navigating the application

Tabs

A tab appears only when the corresponding feature is enabled at your site and you have the appropriate permissions.

If you use one tab most frequently, consider setting your user preferences so that the tab appears after you log in.

Tab	Description
Home	The content of this page can be customized to meet the needs of your company, team, or project. Multiple versions of the Home page might exist, each associated with a different group of users. For more information, see Specifying the Home page for a login group.
Descriptive Analysis	Compute descriptive statistics about specified terms in specified temporal patterns.
Evaluative Analysis	Perform cohort analysis.
Patient Queries	Work with queries, which are used to create patient groups. You use patient groups to restrict an analysis to a particular set of patients.
Reports	Use reports to retrieve source data, view the data in a tabular format, and view graphs of report data.

Commands

The following commands appear in the upper-right corner of all pages.

Command	Usage
Browse Data	Search quickly for patients with specified criteria.
Preferences	Set your user preferences.
Settings	Perform user and system administration tasks, including changing your password.
Feedback	Send feedback, such as comments, suggestions, or feature requests, to an email address that is specified by your site administrator.
Exit	Exit the application.
Help	Open the Help.

Navigating the user interface

Because the appearance of many pages relies on choices made on previous pages, Oracle recommends that you do not use the Back button in the Microsoft Internet Explorer browser. Instead, click the Back link, which appears on many pages, or navigate using the tabs at the top of the page.

Allowing pop-up windows

Some features of the Empirica Healthcare Analysis application, as well as its Help, use pop-up windows. If your pop-up blocking software prevents the windows from opening, you should allow pop-up windows for the web site on which you run the Empirica Healthcare Analysis application.

Related topics

About the Empirica Healthcare Analysis software

Changing your password

You can change your password at any time. You might want to change it periodically for security reasons.

Some sites specify a password expiration period. If your password is due to expire, a message informs you after you log in. If your password expires before you change it, you can change your password the next time you log in.

If you forget your password or are unable to change it, an administrator can change your password for you.

Password requirements

Passwords are case-sensitive; user names are not.

An application administrator determines the required password length and the types of characters required. There might be restrictions on the re-use of old passwords.

To change your password because you are prompted:

1. Log in.

   A page appears, informing you that your password expired.

2. Type your old and new passwords, and click **Change**.

To change your password from within the application:

1. At the top of any page, click **Settings**.
2. Click **Change Password**.

   The Change Password page appears.

3. Type your old and new passwords, and click **Change Password**.

Setting user preferences

A user preference is a setting that customizes part of the application. Your user preferences do not affect other users.

Note: If you have permission to create user profiles, you can also set user preferences for a user profile. You edit a user profile on the Settings page.

1. At the top of any page, click **Preferences**.

   The Set User Preferences page appears.

2. Modify any of the user preferences described in the following table.

3. Click **Save**.

Preference descriptions

Preference	Description	When change takes effect
Dates and Times		
Time Zone	Time zone that is used to display date time information from the server, such as the date an analysis run was created. This preference affects only your view of date times and does not affect the way that source data is displayed. For example, the run creation date on an analysis tab is affected by the preference change. However, source data such as diagnosis date is not affected by the change.	Immediately.
Logging In		
On login, display this page	Page that appears after you log in. The available options depend on your permissions and the way that the application was set up.	The next time you log in.
Show table scrollbars on left	When selected, vertical scrollbars appear on the left side of tables. This setting is useful for wider tables, for which you would have to scroll to the right to see the vertical scrollbar. When not selected, vertical scrollbars appear on the right side.	The next time you log in.
Hierarchies		

Enable Event Hierarchy Browser	When selected, you can select terms from a hierarchy in most activities where you select diagnoses. This feature is available for diagnosis variables that are set up with the appropriate selection type in the data configuration.	Immediately.
Enable Drug Hierarchy Browser	When selected, you can select terms from a hierarchy in most activities where you select drugs. This feature is available for drug variables that are set up with the appropriate selection type in the data configuration.	Immediately.
Downloading Data		
Default download file type	File type that is selected by default when you download tables. This preference does not affect the downloading of patient details.	Immediately.
Viewing Patient Details		
Days to retain visited status for Patient ID links	Number of days that an underlined patient ID that you click to view patient details retains a different color. This preference applies to patient IDs for: • Only the user name of the person who clicked the patient ID. • All data associated with configurations within the same database group. If the value is 0, when a patient ID link is visited, its color change is for only the specific link that was clicked and is not retained beyond the current display. If the value is empty, the default value of 45 is used. This preference affects only data associated with data configurations that are in database groups.	Immediately.
Wrap Wide Tables	When selected, tables on the Patient Details page are wrapped.	Immediately.
Show Table of Contents	When selected, a table of contents appears on the Patient Details page.	Immediately.
Display Patient Timelines as	(Available only when the application is set up to show this option.)	Immediately.

applet	When selected, single-patient and multi-patient timelines appear with Java-enabled interactive features. When deselected, timelines are static JPEG images without interactive features.

Specifying settings

Click **Settings** at the top of any page to specify the following settings. Your user permissions determine the options that are available.

Setting up the application involves creating users, setting site options, and setting up configurations.

- Change Password
- Remove Visited Status for Patient ID Links
- Manage Saved Lists
- About

Manage Users

- Edit Users
- Edit User Profiles
- Edit Roles
- Edit Login Groups

Monitor System

- View Currently Logged In Users
- View User Activity Audit Trail
- View Server Status
- View Free Space

Administer System

- Set Site Options
- Send Message to All Users
- Restart Listener

Configure System

- Manage Configurations

Exiting the application

You should exit the application before closing your internet browser. If you close your browser without exiting the Empirica Healthcare Analysis application, your session continues to run for some time before the automated timeout and might use system resources unnecessarily.

You can exit the application while analysis runs are running in the background or are in a queue to run. Runs continue after you exit.

To exit:

- Click **Exit**, located at the top of any page.

  A message appears, indicating that you are logged out of the application.

Related topics

About the Empirica Healthcare Analysis software

Sending feedback

Oracle welcomes your comments, problem reports, suggestions, or requests about the Empirica Healthcare Analysis application.

1. At the top of any page, click **Feedback**.
2. Fill in the fields, and click **Send**.

   A confirmation appears, informing you that your message was sent.
3. Click **Continue**.

Browsing patient data and creating a patient group

About browsing patient data

For information about creating a patient group, see Browsing patient data and creating a patient group.

Setup and panels

Oracle Consulting Services sets up the Browse Data feature. Typically, the Browse Data dialog box includes panels for the following information:

- Several demographic variables, such as gender and age group, sorted alphabetically.
- A drug variable.
- A diagnosis variable.
- A procedure variable, if procedures are supported by the configuration.

Patient counts

After you select values for variables in the Browse Data dialog box, counts are generated for the patients who meet the specified criteria.

The enrollment dates of patients affect the patient counts in the following ways:

- For diagnosis or procedure variables, patient counts include only diagnoses or procedures that occur within a patient enrollment period.
- For drug variables, patient counts include only prescriptions that start within a patient's enrollment period. The prescriptions do not need to end within the patient's enrollment period.

Logic

The logic between variables is as follows:

- Values selected for a variable in a panel are joined by OR.

  For example, if you select Female and Male for Gender, patients who are either female or male are counted.
- Criteria for different variables within a panel are joined by AND.
- Criteria across panels, either different types or the same types, are joined by AND. Therefore, the total combined count of patients includes patients who meet the criteria of all the panels.

Browsing patient data and creating a patient group

The Browse Data feature allows you to find counts of patients who match specified criteria for demographic variables, drug variables, diagnosis variables, and procedure variables. From the patient counts in the Browse Data dialog box, you can view details about each patient, view single-patient timelines, and create patient groups, and perform other tasks.

The Browse Data feature allows you to create a patient group for which the patient IDs are saved, but the criteria for identifying the patients are not saved.

For more information, see About browsing patient data.

Prerequisites

- To browse data, you must have the Browse Data permission.
- To create a patient group, you must have the Create Queries/Patient Groups permission.

To browse patient data and create a patient group:

1. Click **Browse Data**, located in the row of links in the upper-right corner of the application.

   The Browse Data dialog box appears.

2. Select a configuration from the **Configuration** drop-down list.

   Panels appear for selecting criteria for the search.

3. Select values for the search. You can specify values for any combination of variables. Properties of the variable and the number of its distinct values determine the type of controls you use to select values:

   - If a variable has ten or fewer values, the values appear as buttons.
   - If a variable has more than ten values, a **Select Available Values** link appears for selecting values from a drop-down list.
   - If a variable has an associated hierarchy of terms and your preferences enable the hierarchy, a **Select <*hierarchy name*> Terms** link appears for selecting terms from a hierarchy.
   - For a drug, diagnosis, or procedure variable, the **Select Saved List** link appears for selecting values from a saved list.

   After you specify values, counts for the patients that meet the criteria of each panel and the criteria of all panels appear in the dialog box.

   For information about additional tasks you can perform on this page, see:

   - About drilldown.
   - Viewing a list of patients.

- Creating a patient group.

4. (Available if the panel displays links) To clear selected values for a panel, click **Clear** in the panel.

5. To work with panels:

   - To add a panel, click the **Add Panel** drop-down list located above the panels, and select the type of panel to add.
   - To delete a panel, click **Delete**, located next to the panel name.
   - To collapse a panel, click the negative (-) button to the left of the panel name.
   - To expand a panel, click the positive (+) button to the left of the panel name.

6. After you have finished specifying values and are ready to create a patient group, click **Create Patient Group**, located at the bottom of the dialog box.

Note: Before you can create a patient group, you must specify values for the search, and the application must compute the total counts.

The Create Patient Group dialog box appears.

7. Fill in the fields, and click **Create**.

   The patient group is created, and the dialog box closes.

8. When you are finished creating patient groups, click **Close**.

Descriptive Analysis tab

Descriptive Analysis page

In the Empirica Healthcare Analysis application, you can perform several types of analysis, including descriptive analysis. A descriptive analysis run generates counts and statistics for specified drugs, diagnoses, or procedures during specified periods and in specified temporal patterns that occur in patients' healthcare records. You can search in either the entire source dataset or in selected groups of patients.

You create descriptive analysis runs and view information about runs, including their results, on the Descriptive Analysis tab. The page lists the following runs:

- Runs that you created.
- Runs that are published to you or your login group.

Note: Your view might include runs that are listed for configurations to which you have not been granted permission.

General activities

The following links and filters appear at the top of the page and affect the entire page.

- Create Descriptive Analysis
- Columns
- Print
- Download
- Select Rows
- Filter by Project or Configuration

Row-specific activities

The following menu options are available from the row menu, located in the leftmost column of the table, and affect an individual row in the table. Your permissions determine the options that appear.

- View Results
- Rename
- Cancel (for runs that are in progress) or Delete (for completed runs)
- Re-run
- View Jobs for Run
- View Run Details
- Publish

Field descriptions—Descriptive Analysis page

Field	Description
Configuration	Name of the configuration used by the run.
Created	Date and time when the run was created.
Created By	User who created the run.
Description	Description of the run.
End Date	Date and time when the run ended. This column is empty until all the jobs are complete.
ID	Identifier that was assigned to the analysis run when the run was submitted. The run IDs are unique across the Descriptive Analysis and Evaluative Analysis tabs. An ID of a deleted run is not reused.
Name	Name of the run.
Project	Name of the project to which the run is assigned.
Run Type	One of the following values, indicating the type of descriptive analysis run: • **Risk-Outcome**. • **Drug Utl.**—In the Drug Utilization analysis, a drug variable was used for the outcome event. • **Outcome Char.**—In the Outcome Characterization analysis, a diagnosis or procedure variable was used for the outcome event.
Size	Size of the results table.
Start Date	Date and time when the run started. For a scheduled run, the start date might be substantially later than the creation date of the run.
Status	The column is empty until all jobs are complete. • **Completed** appears after all jobs complete successfully. • **Error Occurred** appears if a job failed, resulting in a failed run. To learn more about the error, select the row menu ( ) for the run, and select **View Jobs for Run**. Then review the Status column. • **Cancelled** appears if the run was cancelled.

Descriptive analysis runs

A descriptive analysis run is a background process that can start immediately or at a future time. An analysis run that starts at a future time is a scheduled run.

You can define an analysis run to break down results according to demographic characteristics, time categories (year, quarter, or month), or co-occurring terms. Co-occurring terms are other drugs, diagnoses, or procedures that occurred within a specified range relative to a reference date.

Co-occurring terms might be concomitant medications, drug history, comorbidities, or medical history. For example, if you define a run to look at the occurrence of liver abnormalities following exposure to statins, you could also look for the ten most frequently occurring diagnoses that occur during the exposure.

Types of descriptive analysis runs

Type of descriptive analysis	What is computed
Risk-Outcome analysis	Incidence rate for an outcome during a user-defined follow-up period that is relative to an index date. The index date is the occurrence of a specified drug, diagnosis, or procedure. The incidence rate is the rate of occurrence of the outcome event per year, according to the following calculation: Occurrence of outcome during follow-up / person-days at risk For example, an analysis could look at occurrence of cardiac events during the first 30 days after a particular drug is administered. The outcome event and index event can be one of the following: • One or more drugs. • One or more diagnoses. • One or more procedures. Depending on how you define follow-up, the follow-up start can be either of the following: • The date of the index. • A specified number of days after the index event. The follow-up end is the earliest of the following possibilities: • Occurrence of the outcome. • Specified duration of follow-up (optionally specified). • End of index exposure (optionally specified if the index is a drug). • End of enrollment for patient.

	• End of the specified analysis period.
Drug Utilization analysis	Period prevalence for exposure to a drug. Period prevalence represents patients who were exposed to the drug out of all patients included in the analysis.
Outcome Characterization analysis	Period prevalence for the occurrence of a diagnosis or procedure. Period prevalence represents patients who had the diagnosis or procedure out of all patients included in the analysis.

Analysis results

A descriptive analysis run looks at data for each patient included in the analysis and applies the parameters of the run to that data. You specify an analysis period. Only patients who have any enrollment days within the specified analysis period are included in the analysis.

- For a Drug Utilization or Outcome Characterization analysis, the application generates counts of patients who are included in the analysis and have specified terms.
- For a Risk-Outcome analysis, the application generates the following information:
  - Counts of patients who are included in the analysis and have specified terms within a defined follow-up period, also known as the risk window.
  - Counts of person-years in follow-up; that is, a count of days in follow-up, converted into years.

The analysis uses the counts to compute statistics, such as period prevalence and incidence rate.

In some cases, the analysis also computes counts of diagnoses, procedures, or prescriptions.

Enrollment

A patient's enrollment period is the period of time during which the patient could have contributed data to the healthcare database, as determined by the configuration that you are using. A configuration is a group of variables representing items (database table columns) in the source data.

When a configuration is set up, a decision is made about which variables in the source data will be used to indicate the start and end of patient enrollment periods. For example, suppose that the source data is insurance claims data. The start and end dates of eligibility for the insurance plan could be used as **enrollment start** and **enrollment end**. If patients are eligible during multiple periods of time because they leave the health plan and join it again, the enrollment period could be set up as the longest enrollment period or it could be set up as the start of the first eligibility period and the end of the last.

Temporal patterns

In generating counts, the analysis looks at temporal patterns. To determine whether temporal criteria are met, the analysis characterizes each patient day as occurring before, during, or after various points or periods of interest. The points or periods of interests are

defined by the values of the temporal parameters you specify and by the days on which specific prescriptions start or specific diagnoses or procedures occur.

Consider, for example, a Risk-Outcome analysis run that uses an index lead-in of ten days, an outcome lead-in of 30 days, and a follow-up that begins three days after the index date and ends 60 days later. The analysis establishes the following windows around each potential index date, referred to in the following examples as day 0:

- An index event wash-out window runs from day -10 to day 0.
- A dosing window runs from day 0 to day 3.
- An outcome event wash-out window runs from day -27 to day 3.
- A risk window runs from day 3 to day 63.

The presence or absence of index and outcome events within these various windows serve to establish the index and incidence dates for each patient.

You can establish additional windows of interest through the use of the following options:

- Breakdowns based on time categories (year, quarter, and month).
- Co-occurrence windows, which identify relevant concomitant medications and/or medical history.

Types of data

You can specify drugs, diagnoses, or procedures for the following items in a descriptive analysis run:

- Index event
- Outcome event
- Co-occurrence term

Descriptive analysis runs ignore the following information:

- Prescriptions that start outside of a patient's enrollment period.
- Diagnoses or procedures that occur outside of a patient's enrollment period.

Diagnoses and procedures

A diagnosis is recorded in the source data for a patient. For example, the diagnosis from insurance claims data might be a term from the ICD-9-CM.

A diagnosis or procedure is a time point consisting of a single day. Diagnoses and procedures are determined from the following information in the source data:

- Diagnosis term or procedure term.
- Date that represents the occurrence of the diagnosis or procedure.

If you select multiple diagnoses or multiple procedures as the index event or outcome event in a run, the analysis treats the values as if they are one term and looks for an occurrence of any values.

Prescriptions

A prescription is a time interval consisting of one or more days. Prescriptions are determined from the following information in the source data:

- Drug term.
- Date that represents the prescription start.

  For example, the prescription's dispensation date or, for mail-order prescriptions, the dispensation date plus two days.
- Days supplied.

The prescription start date is the first day of the prescription. For example, consider a prescription that starts on March 1, 2008 and is supplied for four days. The first date of the prescription is March 1, 2008, and the last day is March 4, 2008.

If you select multiple drugs as the index event or outcome event in a run, the analysis treats the values as if they are one term.

Drug exposures and drug eras

During analysis runs, drug eras are generated from prescriptions. A drug era is an interval of time during which a patient is considered to be continuously exposed to a specified drug or group of drugs. When analyzing drug information, an analysis run looks at either prescriptions or drug eras, depending on the context.

Note: Prescription data is from the source data. Drug eras are constructed from the source data during analysis runs but do not change the source data.

Prescriptions that start before the patient's enrollment start are ignored even if they continue into the patient's enrollment period. For example, consider the following situation:

- A ten-day prescription starts on June 1, 2006.
- The patient's enrollment start is June 5, 2006.
- A three-day prescription starts on June 8, 2006.

The first drug era is from June 8, 2006 through June 10, 2006.

Drug eras are constructed from the source data as follows:

1. The end of each prescription is determined.

   For example, suppose that a prescription starts on June 1, 2006 and the days supplied is 10. The prescription ends on June 10, 2006.

2. If two or more prescriptions are contiguous or overlapping, they are combined into a single drug exposure.

3. If no exposure extension is specified, each resulting drug exposure is a drug era.

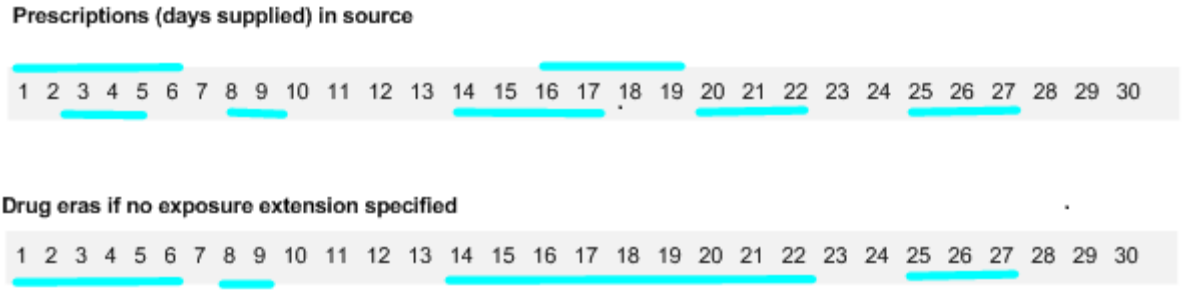
   If you specify an exposure extension (a number of days), the extension is applied as follows:

A. The specified number of days is added to the end of each resulting drug exposure identified in the first step.

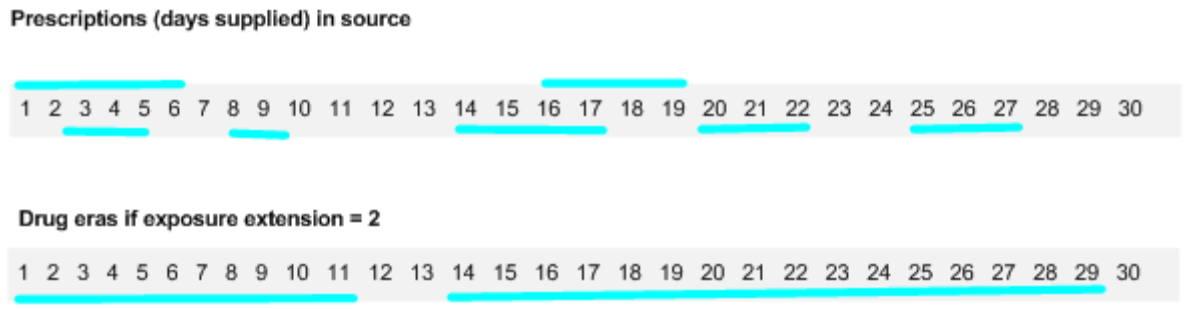
B. If the number of days between the end of one prescription and the start of the next prescription is less than or equal to the specified number of days, the prescriptions are combined. If the number of days between the two prescriptions is greater than the specified number of days, the prescriptions are not combined.

The result is one or more drug eras.

In the following example, four drug eras are constructed from the patient's prescriptions in the source data:



The following example shows how prescriptions are combined if an exposure extension is specified. Two drug eras are constructed.



Creating a Risk-Outcome descriptive analysis run

Step 1—Creating a descriptive analysis run

Prerequisites

- The application must have been set up so the Descriptive Analysis tab is visible.
- To create a run, you must have the Create Analysis Run permission.
- To view results of a run, you must have the View Analysis Results permission.
- Either you or your login group must have permissions for the configuration.

To create a descriptive analysis run:

1. Select the **Descriptive Analysis** tab.
2. Click **Create Descriptive Analysis**.

   The Create Descriptive Analysis page appears.
3. Fill in the fields according to the following table, and click **Next**.
4. Continue to one of the following steps:
   - For a Risk-Outcome analysis—Step 2a—Defining an index event.
   - For a Drug Utilization or Outcome Characterization analysis—Step 2—Defining an outcome event (Drug Utilization or Outcome Characterization).

Field descriptions—Create Descriptive Analysis page

Field	Description

Select the type of analysis to create	• **Risk-Outcome**—Create a Risk-Outcome analysis run. You specify:
	  • An index event based on a drug, diagnosis, or procedure variable.
	  • An outcome event based on a drug, diagnosis, or procedure variable.
	• **Drug utilization / Outcome characterization**—Create either a Drug Utilization or an Outcome Characterization analysis run. For both analysis runs, you specify an outcome event. If the outcome event is based on:
	  • A drug variable—The analysis is a Drug Utilization analysis run.
	  • A diagnosis or procedure variable—The analysis is an Outcome Characterization analysis run.

Configuration	Configuration to use for the analysis run. Select a configuration by choosing it from the drop-down list or by

	clicking **Browse**.
Choose initial cohort	Patients against which the analysis is run. • **Use existing patient group**—Run the analysis against the patient group that you select from the drop-down list. You can also click **Browse** to navigate to a patient group. To create a patient group, click **Browse Data** at the top of the page. The new patient group that you create is added to the **Use existing patient** drop-down list immediately. For more information, see Browsing data. • **Use all patients**—Run the analysis against all patients in the selected configuration. During the analysis, patients whose enrollment period does not overlap with the specified analysis period are dropped. For example, the following patient who is included in the initial cohort is dropped from the analysis and is not represented in the N(Initial) column of analysis results: • Enrollment period—January 1, 2006 - January 1, 2007. • Analysis period start—May 1, 2007.
Start of analysis period	(Available after you select an initial cohort.) First day of data to analyze for the run. By default, the earliest enrollment start date for any patient in the initial cohort (either a patient group or all patients) appears. You can select a date that is later than the default but not earlier.
End of analysis period	(Available after you select an initial cohort.) Last day of data to analyze for the run. By default, the latest enrollment end date for any patient in the initial cohort (either a patient group or all patients) appears. You can select a date that is earlier than the default but not later.

Step 2a—Defining an index event

Previous step	Step 1—Creating a descriptive analysis run.
Relevance	Only if you are creating a Risk-Outcome analysis run.
Description	In this step, you define an index event, which is a list of values that the analysis run uses to establish an index date for each patient. Defining an index event consists of the following steps: 1. Selecting a drug, diagnosis, or procedure variable. 2. Selecting values for the variable. Later in the wizard, you specify temporal parameter criteria. The analysis attempts to determine an index date for each patient using the specified index event and the temporal parameter criteria.

To define an index event:

1. On the Define Index Event page, fill in the fields according to the following table, and click **Next**.

   The Define Outcome Event page appears.

2. Continue to Step 2b—Defining an outcome event.

Option descriptions—Define Index Event page

Option	Description
Show Preliminary Counts	View preliminary counts of patients to help determine if you are selecting appropriate values. The preliminary counts section dynamically displays counts of possibly eligible patients based on the parameters selected for the analysis run. The counts change as you select values on each page that displays the Preliminary Counts section, allowing you to determine whether the preliminary set of patients potentially will yield meaningful outcome data before you perform the analysis run. The preliminary counts section displays more information as you proceed through the wizard. For each cohort, they can include potential patients with the index term, the outcome term, and incidence (both the index term and the outcome term). The preliminary counts do not use the specified analysis period or any temporal parameters of the run. Therefore, a patient counted in preliminary counts in the run wizard might not be in counts in the actual analysis results.
Select index variable	Index event variable. Available variables are determined by the configuration you are using.

Define index event as _____ of terms selected below	Index type. For a diagnosis or procedure variable, the following types are available: • start of first occurrence • start of last occurrence For a drug variable, the following types are available: • start of first occurrence • start of last occurrence • end of last occurrence For information about each of these options, see Determination of the index date.
If selecting more than one value, provide name for custom term	Name for the custom term defined as the index event. If you select multiple values in the list box, you must provide a name.
[List box for selecting values]	See Selecting values.
Remove terms if they are not in the source data	When selected, terms that are in the list of specified values and that are not in the source data are removed from the list when you continue to the next page of the wizard.
Exposure extension: ___ days	(Available for a drug variable.) When selected, you can specify an exposure extension in days.

Step 2b—Defining an outcome event (Risk-Outcome)

Previous step	Step 2a—Defining an index event.
Relevance	Only if you are creating a Risk-Outcome analysis run
Description	In this step, you specify one or more values as the outcome event, which can be based on a drug, diagnosis, or procedure variable. The analysis attempts to find the occurrence of the outcome event during follow-up. • If the outcome event is based on a diagnosis or procedure variable, the analysis looks for the first occurrence of the outcome event during follow-up. • If the outcome event is based on a drug variable, the analysis looks for the first day of the first drug era that starts during follow-up.

To define an outcome event in a Risk-Outcome analysis:

1. On the Define Outcome Event page, fill in the fields according to the following table, and click **Next**.

   The Descriptive Analysis: Specify Temporal Parameters page appears.

2. Continue to Step 2c—Specifying temporal parameters.

Option descriptions—Define Outcome Event page

Option	Description
Show Preliminary Counts	View preliminary counts of patients to help determine if you are selecting appropriate values. The preliminary counts section dynamically displays counts of possibly eligible patients based on the parameters selected for the analysis run. The counts change as you select values on each page that displays the Preliminary Counts section, allowing you to determine whether the preliminary set of patients potentially will yield meaningful outcome data before you perform the analysis run. The preliminary counts section displays more information as you proceed through the wizard. For each cohort, they can include potential patients with the index term, the outcome term, and incidence (both the index term and the outcome term). The preliminary counts do not use the specified analysis period or any temporal parameters of the run. Therefore, a patient counted in preliminary counts in the run wizard might not be in counts in the actual analysis results.
Select outcome variable	Outcome event variable. Available variables are determined by the configuration you are using.

If selecting more than one value, provide name for custom term	Name for the custom term defined as the outcome event. If you select multiple values in the list box, you must provide a name.
[List box for selecting values]	See Selecting values.
Remove terms if they are not in the source data	When selected, terms that are in the list of specified values and that are not in the source data are removed from the list when you continue to the next page of the wizard.
Exposure extension	(Available for a drug variable.) When selected, you can specify an exposure extension in days.

Step 2c—Specifying temporal parameters

Previous step	Step 2b—Defining an outcome event (Risk-Outcome).
Relevance	Only if you are creating a Risk-Outcome analysis run.
Description	In this step, you can specify: • An index lead-in • An outcome lead-in. • A period of follow-up, also known as a risk window, that is always relative to the index date. If you specify an index lead-in or an outcome lead-in, the analysis uses the values when attempting to establish an index date for a patient. If the lead-in criteria are not met, an index date is not established. To determine whether the outcome lead-in is met, the analysis must determine how the start of follow-up is specified. Therefore, the options on this page are used together to determine whether a patient has an index date and, if so, what the follow-up is.

To specify temporal parameters:

1. On the Descriptive Analysis: Specify Temporal Parameters page, fill in the fields according to the following table, and click **Next**.

   The Descriptive Analysis: Define Breakdowns for Results page appears.

2. Continue to Step 3—Defining breakdowns for results.

Option descriptions—Descriptive Analysis: Specify Temporal Parameters page

Option	Description
Show Preliminary Counts	View preliminary counts of patients to help determine if you are selecting appropriate values. The preliminary counts section dynamically displays counts of possibly eligible patients based on the parameters selected for the analysis run. The counts change as you select values on each page that displays the Preliminary Counts section, allowing you to determine whether the preliminary set of patients potentially will yield meaningful outcome data before you perform the analysis run. The preliminary counts section displays more information as you proceed through the wizard. For each cohort, they can include potential patients with the index term, the outcome term, and incidence (both the index term and the outcome term). The preliminary counts do not use the specified analysis period or any temporal parameters of the run. Therefore, a patient counted in preliminary counts in the run wizard might not be in counts in the actual

	analysis results.
No occurrence of index term *<index term>* for ___ days before index date	(Available when the index date is based on the first occurrence of a diagnosis, procedure, or drug. Not available when the index event is based on the start or end of the last occurrence.) When selected, you can specify an index lead-in, which is the number of days before the index date that must be clear of the value specified as the index event. If the index event is a drug, the days must be clear of any part of a drug era for the drug. The clear days: • Occur immediately before the index event and are counted backwards from the index event, treating the index date as Day 0. • Occur within the patient's enrollment period, but they do not need to be within the analysis period specified for the run. For example, if the analysis period starts 100 days after the patient's enrollment start, the days can be within those 100 days before the analysis period start. The *<index term>* value in the label is one of the following: • The single value you specified. If the value is too long to fit, you can point to the underlined value to see a tooltip. • The custom term you specified. Point to the underlined custom term to see its values. If you do not specify an index lead-in: • There is no requirement for a clear period before the index date. • You cannot specify a co-occurrence lead-in.
No occurrence of *<outcome term>* outcome term for ___ days before follow-up start	When selected, you can specify an outcome lead-in, which is the number of days before follow-up start that must be clear of the value specified as the outcome event. If the outcome event is a drug, the days must be clear of any part of a drug era for the drug. The clear days: • Occur immediately before the follow-up start and are counted backwards from the follow-up start, treating the index date as Day 0. These days are before follow-up start, which could start after the index date. • Occur within the patient's enrollment period, but they do not need to be within the analysis period specified for the run. For example, if the analysis period starts 100 days after the patient's enrollment start, the days can be within those 100 days before the analysis period start. The *<outcome term>* value in the label is one of the following: • The single value you specified. If the value is too long to fit, you can point to the underlined value to see a tooltip. • The custom term you specified. Point to the underlined custom term to see its values.

	If you do not specify an outcome lead-in, there is no requirement for a clear period before the follow-up start.
Specify start of follow-up (risk window) as	• **Date of index event**—When selected, follow-up starts on and includes the index date. • **Follow-up begins ___ days after index event**—When selected, follow-up starts on the specified number of days after the index date. The index date is day 0, so a value of 1 is the day after the index date. For example, when a patient's index date is June 1, 2006 and you specify a follow-up start of 5 days after the index event, follow-up starts on June 6, 2006. This option does not preclude the possibility of an outcome event occurring after the index and before the follow-up start.
Specify end of follow-up (risk window) as earliest of these options Follow-up ends for a patient when any of the following selected options occurs.	
after ___ days of follow-up	When selected, follow-up ends after the specified number of days. For example, if you specify follow-up start as the index date, with an index date of June 1, 2006, and you specify 5 for this value, the last day of follow-up is June 5, 2006. If you specify follow-up start as 5 days after the index event, the first day of follow-up is June 6, 2006. If you specify 6 for the **after ___ days of follow-up** value, the follow-up end is June 11, 2006.
end of index exposure period	(Available when the index is based on the start of the first occurrence of a drug or the start of the last occurrence of a drug.) When selected: • If the index date is the start of the first occurrence of a drug exposure, follow-up ends on the last day of the drug era. • If the index date is the start of the last occurrence of a drug exposure, follow-up ends on the last day of the drug era that encompasses the prescription that determines the index date. Drug eras are used in determining the end of index exposure period and thus the end of follow-up, even though they are not used to determine the index date itself. If the follow-up start is specified in a way that would make follow-up start after the end of the drug era for a candidate index exposure, the index candidate is rejected. For example, if the drug era for the candidate index is from May 1, 2005 to May 5, 2005, and the follow-up start is specified as 5 or more, the index candidate is rejected. The index date is counted as day 0, so 5 days after index is May 6, 2005.
upon first occurrence of outcome event	(Always selected; cannot be unselected.) • If the outcome event is a diagnosis or procedure, follow-up ends on the date of the first occurrence of the outcome event during follow-

	up. • If the outcome event is a drug, follow-up ends on the first day of the first drug era that starts during follow-up. If the outcome event is based on a drug variable and a drug era starts before follow-up and continues into follow-up, the patient is not counted as having the outcome event during follow-up unless the patient has a separate, subsequent drug era that starts within follow-up, regardless of whether an exposure extension is used.
upon end of analysis period or enrollment	(Always selected; cannot be unselected.) Follow-up ends at the end of the patient's enrollment period or the end of the specified analysis period for the analysis run.

Step 3—Defining breakdowns for results

Previous step	One of the following: • Step 2c—Specifying temporal parameters. • Step 2—Defining an outcome event (Drug Utilization or Outcome Characterization).
Relevance	All types of descriptive analysis runs.
Description	In this step, you define breakdowns of analysis results. The results table contains one row for each breakdown. The breakdown value is in the SUBSET column of results.

To define breakdowns of analysis results:

1. On the Descriptive Analysis: Define Breakdowns for Results page, fill in the fields according to the following table, and click **Next**.
2. One of the following pages appears:
   - If you selected One or more co-occurrence windows, the Specify Parameters for Co-Occurrence page appears. Continue to Step 4—Specifying parameters for a co-occurrence window.
   - If you did not select **One or more co-occurrence windows** and you have the Access Advanced Run Options permission, the Descriptive Analysis: Advanced Options page appears. Continue to Step 5—Specifying advanced options.
   - If you did not select **One or more co-occurrence windows** and you do not have the Access Advanced Run Options permission, the Run Options page appears. Continue to Step 6—Specifying run options.

Field descriptions—Descriptive Analysis: Define Breakdowns for Results page

Field	Description
Break down the analysis results by the following	
<Demographic variables>	Demographic variables as determined by the configuration. For example, Gender is typically available as a breakdown. In the analysis results, one subset row appears for each value of the demographic variable that any patients in the initial cohort have. For example, when a patient group of only males is used as the initial cohort and the results are broken down by gender, there is only an ALL subset and a subset for males. In the SUBSET column of analysis results, the breakdown appears as ***<column name>*:** followed by the variable value for that breakdown, such as **GENDER: F**. If the value is not in the source data, the breakdown value is **null**.

Custom Age Group (of computed age)	(Appears if the configuration is set up to support custom age groups.) The computed age group is based on a computed age and on age groups that you define. In analysis results, one subset row appears for each computed age group that you define. Patients are in the age group if their computed age is within that range of the age group. In the SUBSET column of analysis results, the breakdown appears as **AGEGRP (Custom):** followed by the computed age group. Notes: • If the age group cannot be computed, the breakdown appears as **AGEGRP (Custom): Unknown**. • If the computed age is a negative number, the breakdown appears as **AGEGRP (Custom): Less than 0**. For more information, see Working with age groups.
Time category	Break down results for the following time categories: • **Year**—Create a breakdown for each calendar year, starting in January, within the analysis period defined for the analysis. In the analysis results, the subset names are **YEAR: *<year>***. • **Quarter**—Create a breakdown for each quarter, starting in January, within the analysis period defined for the analysis. In the analysis results, the subset names are **QUARTER: *<year and quarter>***. • **Month**—Create a breakdown for each month within the analysis period defined for the analysis. In the analysis results, the subset names are **MONTH: *<year and month>***.
One or more co-occurrence windows	When selected, the Specify Parameters for Co-Occurrence Window page appears later in the wizard, allowing you to identify co-occurring values during a specified time period. For more information, see Step 4—Specifying parameters for a co-occurrence window.

Step 4—Specifying parameters for a co-occurrence window

Previous step	Step 3—Defining breakdowns for results.
Relevance	Any type of descriptive analysis, as long as you selected **One or more co-occurrence windows** when defining breakdowns for results.
Description	In this step, you decide whether to identify co-occurring values, such as co-morbidities, medical history, and concomitant medications, during a specified time period, which is the co-occurrence window. You can define multiple co-occurrence windows. A co-occurrence window is a type of breakdown of results, resulting in subset rows for co-occurrence windows in the analysis results. For a co-occurrence value to appear as a row in results, at least one patient must have the co-occurring value. To define a co-occurrence window, you perform the following tasks: • Define a time period relative to a reference date. • Select a variable to be used for the co-occurrence window. • Either explicitly select co-occurring values to look for, or request breakdowns by the most frequent values within the co-occurrence window. If the co-occurrence window is based on a: • Diagnosis or procedure variable—Diagnoses or procedures that occurred within the window are considered to be co-occurring. • Drug variable—Drugs with any part of a drug era during the window are considered to be co-occurring.

To specify parameters for a co-occurrence window:

1. On the Specify Parameters for Co-occurrence Window page, fill in the fields according to the following table, and click **Next**.
2. To delete the co-occurrence window, click **Delete**, located at the bottom of the page, near the Back and Next buttons.

   A confirmation message appears. If you confirm the deletion, the current window is deleted. If you delete all co-occurrence windows, the **One or more co-occurrence windows** checkbox on the Define Breakdowns for Results page is deselected.
3. To add another co-occurrence window, click **Add**, located at the bottom of the page, near the Back and Next buttons.

   The window is added immediately after the current window. The current window is saved and a new, blank Specify Parameters for Co-occurrence Window page appears. You can define an unlimited number of co-occurrence windows.

4. Click **Next**.

One of the following pages appears:

- If you added another co-occurrence window, the Specify Parameters for Co-occurrence Window page appears. Fill in the fields according to the following table.
- If you did not add a co-occurrence window:
  - If you have the Access Advanced Run Options permission, the Descriptive Analysis: Advanced Options page appears. Continue to Step 5—Specifying advanced options.
  - If you do not have the Access Advanced Run Options permission, the Run Options page appears. Continue to Step 6—Specifying run options.

Field descriptions—Specify Parameters for Co-occurrence Window page

Field	Description
Window name (prefix)	Name of the co-occurrence window. The values that you select later on the page determine the text that appears in the SUBSET column in the analysis results. If you: • Select values to include or exclude, the column displays the window name. • Select **Enumerate ___ most frequently occurring values**, the following text appears: ***<window-name>: <frequent value>***.
Window description	Description of the co-occurrence window.
Select window variable	Co-occurrence variable. Available variables are determined by the configuration you are using.
Define Co-occurrence Window	Window type. If necessary, you also specify a window start and stop. For more information, see Window types.
Enumerate ___ most frequently occurring values	When selected, co-occurring values are defined as the most frequent values within the co-occurrence window. Type a positive integer that is no greater than 50. For more information, see Most frequent values.
Select specific values Specify one or more inclusion values Specify zero or more	When selected, you can create a co-occurrence window to find the following information: • Patients who have any of up to 25 inclusion values during the co-occurrence window.

exclusion values

- Patients who do not have any of up 25 exclusion values during the co-occurrence window.
- Patients who have any of up to 25 inclusion values during the window and do not have any of up to 25 exclusion values during the window.

You specify the values of the co-occurrence variables.

The selected values are treated as a group for which there is a single breakdown. The name of the breakdown in the SUBSET column of results is ***<window-name>***.

Inclusion values

A patient is considered to have the co-occurrence breakdown if any of the inclusion values are in the co-occurrence window. For example, if you select DrugA, DrugB, and DrugC as co-occurring values and name the co-occurrence window ABC, the analysis results contain a single row for which SUBSET is ABC, representing patients who had any of DrugA, DrugB, or DrugC during the co-occurrence window.

Note: If you select the same value that you used to define the outcome event or index event, the value appears as a co-occurring value.

Exclusion values

You can also select exclusion values. If you select only exclusion values, a patient is considered to have the co-occurrence breakdown only if none of the exclusion values are experienced by the patient in the co-occurrence window. For example, if the exclusion values are DrugD, DrugE, and DrugC, the patient cannot have any of those drugs in the window.

Both inclusion and exclusion values

If you specify both inclusion values and exclusion values, a patient has the co-occurrence breakdown only if the following criteria are met:

- The patient has any of the inclusion values in the window.
- The patient does not have any of the exclusion values in the window.

For example, if you specify DrugA, DrugB, and Drug C as inclusion values and DrugD, DrugE, and DrugF as exclusion values, a patient has the co-occurrence breakdown if the patient has any of the following circumstances:

- DrugA during the co-occurrence window and no DrugD, DrugE, or DrugF during the co-occurrence window.
- DrugB during the co-occurrence window and no DrugD, DrugE, or DrugF during the co-occurrence window.
- DrugC during the co-occurrence window and no DrugD,

	DrugE, or DrugF during the co-occurrence window. You cannot specify the same value as both an included value and an excluded value. **Selecting values** You can specify no more than 25 values. For information about selecting values, see Selecting values.
Remove terms if they are not in the source data	When selected, terms that are in the list of specified values and that are not in the source data are removed from the list when you continue to the next page of the wizard.
Exposure extension: ___ days	(Available for a drug variable.) When selected, you can specify an exposure extension in days.
No occurrence of co-occurrence term for ____ days before window start	(Available for a Risk-Outcome analysis run if an index lead-in is specified.) When selected, you can specify a co-occurrence lead-in. The co-occurrence lead-in indicates that the co-occurrence window is immediately preceded by at least the specified number of days clear of the co-occurrence value and within the patient's enrollment period. If the co-occurrence lead-in is not met, a co-occurrence window is not created. The clear days occur within the patient's enrollment period, but they do not need to be within the analysis period specified for the run. For example, if the analysis period starts 100 days after the patient's enrollment start, the days can be within those 100 days before the start of the analysis period. If co-occurrence values are based on a drug variable, this option uses the specified exposure extension option when looking for days clear of the drug. For explicitly included values, none of the selected values can have occurred for the specified number of days before the co-occurrence window. For example, suppose that you specify DrugA, DrugB, and DrugC for inclusion and specify a co-occurrence lead-in of 10. The co-occurrence window must be preceded by at least 10 days in which there are no drug era days for DrugA, DrugB, or DrugC. For explicitly excluded values, none of the selected values can have occurred for the specified number of days before the co-occurrence window. For example, suppose that you specify DrugD, DrugE, and DrugF for exclusion and specify a co-occurrence lead-in of 10. The co-occurrence window must be preceded by at least 10 days in which there are no drug era days for DrugD, DrugE, or DrugF. If you specify DrugA, DrugB, and DrugC for inclusion and DrugD, DrugE, and DrugF for exclusion, and specify a co-occurrence lead-in of 10, the co-occurrence window must be preceded by at least

10 days in which there are no drug era days for DrugA, DrugB, DrugC, DrugD, DrugE, or DrugF. Then, for the patient to have the co-occurrence breakdown, the co-occurrence window must include drug era days for any of DrugA, DrugB, or DrugC and must not include drug era days for any of DrugD, DrugE, or DrugF.

For most frequent values, a patient will not count as having a co-occurrence value unless the co-occurrence lead-in is met. For example, suppose 100 patients have DrugA within what would be the co-occurrence window (if the co-occurrence lead-in was met), but only 10 of the patients meet the co-occurrence lead-in. For the purposes of ranking the most frequently occurring values, DrugA is attributed with having a frequency of 10.

Day ranges

Some co-occurrence window options are for a specified day range relative to a reference date as follows:

- A positive integer indicates days after the reference date.
- A negative integer indicates days before the reference date.
- 0 indicates the reference date itself.

You can indicate a one-day co-occurrence window by entering the same number as the start and end of the window.

Window start and end

When the analysis runs, the start of a co-occurrence window is the later of the following options:

- Patient's enrollment start
- Specified window start

The co-occurrence window can include days that are before the analysis period start and after the patient's enrollment start.

If the specified window start is later than either the patient enrollment end or the analysis period end, a co-occurrence window is not created.

When the analysis runs, the end of a co-occurrence window is the earliest of the following options:

- Enrollment end
- Analysis period end
- Specified window end

Restrictions

Several restrictions ensure that, for a Risk-Outcome analysis, you cannot define a window that would start before a patient's enrollment start.

For a Risk-Outcome analysis, the values you enter must meet the

following criteria:

- If the co-occurrence window is based on the entire index exposure—The co-occurrence lead-in cannot be greater than the index lead-in.
- If the co-occurrence window is based on fixed start/stop relative to outcome date—The co-occurrence lead-in minus the co-occurrence window start cannot be greater than the greater of the index lead-in and the outcome lead-in.
- If the co-occurrence window is based on any other option—The co-occurrence lead-in minus the co-occurrence window start cannot be greater than the index lead-in.

If you do not specify an index lead-in, outcome lead-in, or co-occurrence lead-in, the value of each is considered to be 0. For example, if you do not specify an index lead-in and the co-occurrence window is based on index date, you cannot enter a window start of less than 0.

Step 5—Specifying advanced options

Previous step	One of the following: • Step 3—Defining breakdowns for results. • Step 4—Specify parameters for a co-occurrence window.
Relevance	All types of descriptive analysis runs, but only if you have the Access Advanced Run Options permission.

To specify advanced options:

1. On the Descriptive Analysis: Advanced Options page, fill in the fields according to the following table, and click **Next**.

   The Run Options page appears.

2. Continue to Step 6—Specifying run options.

Field descriptions—Descriptive Analysis: Advanced Options page

Field	Description
Maximum memory available to metrics-generation engine	Maximum amount of memory to devote to the engine that generates metrics. The initial value is the lesser of following values: • The estimate of the memory required for the run. The estimate is fixed at 8.0 GB. • The maximum allowed value as determined during setup.

Step 6—Specifying run options

Previous step	One of the following: • Step 3—Defining breakdowns for results. • Step 4—Specifying parameters for a co-occurrence window. • Step 5—Specifying advanced options.
Relevance	All types of descriptive analysis runs.

To specify run options:

1. On the Run Options page, fill in the fields according to the following table, and click **Next**.

   The Name Analysis Run page appears.

2. Continue to Step 7—Naming an analysis run.

Field descriptions—Run Options page

Field	Description
Run as soon as possible Do not run until ___ at ___	When to execute the descriptive analysis run. Select **Run as soon as possible** to execute the run immediately, or specify the date and time to create the run.
Email me when complete	When selected, an email message is sent to the email address you provide after the run either completes successfully or fails. Separate multiple email addresses with commas.
Save intermediate data files	(Available if you have the Save Intermediate Files permission.) When selected, supplemental data files that are produced when a run is created are saved for viewing and downloading.

Step 7—Naming an analysis run

Previous step	Step 6—Specifying run options.
Relevance	All types of descriptive analysis runs.

To name the analysis run:

1. On the Name Analysis Run page, fill in the fields according to the following table, and click **Next**.

   The Confirm Analysis Run Parameters page appears.

2. Review the selections for the run, and click **Submit Run**.

   A confirmation page appears, indicating that the run has been submitted. You can check the status of the run in the Status column.

   For information about results, see Viewing the results of a descriptive analysis run.

Field descriptions—Name Analysis Run page

Field	Description
Run name Description	Name and description of the run.
Add to	**Existing project**—The run is added to the project selected in the drop-down list. **New project named**—A new project is created with the name you provide, and the run is added to it.

Creating a Drug Utilization or Outcome Characterization descriptive analysis run

Step 1—Creating a descriptive analysis run

Prerequisites

- The application must have been set up so the Descriptive Analysis tab is visible.
- To create a run, you must have the Create Analysis Run permission.
- To view results of a run, you must have the View Analysis Results permission.
- Either you or your login group must have permissions for the configuration.

To create a descriptive analysis run:

1. Select the **Descriptive Analysis** tab.
2. Click **Create Descriptive Analysis**.

   The Create Descriptive Analysis page appears.
3. Fill in the fields according to the following table, and click **Next**.
4. Continue to one of the following steps:
   - For a Risk-Outcome analysis—Step 2a—Defining an index event.
   - For a Drug Utilization or Outcome Characterization analysis—Step 2—Defining an outcome event (Drug Utilization or Outcome Characterization).

Field descriptions—Create Descriptive Analysis page

Field	Description

Select the type of analysis to create	• **Risk-Outcome**—Create a Risk-Outcome analysis run. You specify:
	  • An index event based on a drug, diagnosis, or procedure variable.
	  • An outcome event based on a drug, diagnosis, or procedure variable.
	• **Drug utilization / Outcome characterization**—Create either a Drug Utilization or an Outcome Characterization analysis run. For both analysis runs, you specify an outcome event. If the outcome event is based on:
	  • A drug variable—The analysis is a Drug Utilization analysis run.
	  • A diagnosis or procedure variable—The analysis is an Outcome Characterization analysis run.

Configuration	Configuration to use for the analysis run.

	Select a configuration by choosing it from the drop-down list or by clicking **Browse**.
Choose initial cohort	Patients against which the analysis is run. • **Use existing patient group**—Run the analysis against the patient group that you select from the drop-down list. You can also click **Browse** to navigate to a patient group. To create a patient group, click **Browse Data** at the top of the page. The new patient group that you create is added to the **Use existing patient** drop-down list immediately. For more information, see Browsing data. • **Use all patients**—Run the analysis against all patients in the selected configuration. During the analysis, patients whose enrollment period does not overlap with the specified analysis period are dropped. For example, the following patient who is included in the initial cohort is dropped from the analysis and is not represented in the N(Initial) column of analysis results: • Enrollment period—January 1, 2006 - January 1, 2007. • Analysis period start—May 1, 2007.
Start of analysis period	(Available after you select an initial cohort.) First day of data to analyze for the run. By default, the earliest enrollment start date for any patient in the initial cohort (either a patient group or all patients) appears. You can select a date that is later than the default but not earlier.
End of analysis period	(Available after you select an initial cohort.) Last day of data to analyze for the run. By default, the latest enrollment end date for any patient in the initial cohort (either a patient group or all patients) appears. You can select a date that is earlier than the default but not later.

Step 2—Defining an outcome event (Drug Utilization or Outcome Characterization)

Previous step	Step 1—Creating a descriptive analysis run.
Relevance	Only if you are creating a Drug Utilization or Outcome Characterization analysis run.
Description	For both types of run, you specify an outcome event but no index event. The application needs an index date so that it can break down results. Therefore, the application determines an index date based on the values specified as the outcome event. • For an Outcome Characterization analysis, the outcome event is a diagnosis or procedure. The index date for a patient is the first occurrence of the diagnosis or procedure during the enrollment period and the specified analysis period. • For a Drug Utilization analysis, the outcome event is a drug. The index date for a patient is the first day of a drug era during the enrollment period and the analysis period. If an exposure starts before the analysis period but within the enrollment period and extends into the analysis period, the index date for a Drug Utilization analysis is the same as the analysis period start date. This rule applies even if there is a subsequent and separate drug era that starts during the patient's enrollment period and the specified analysis period.

To define an outcome event:

1. On the Define Outcome Event page, fill in the fields according to the following table, and click **Next**.

   The Define Descriptive Analysis: Define Breakdowns for Results page Event page appears.

2. Continue to Step 3—Defining breakdowns for results.

Option descriptions—Define Outcome Event page

Option	Description
Show Preliminary Counts	View preliminary counts of patients to help determine if you are selecting appropriate values. The preliminary counts section dynamically displays counts of possibly eligible patients based on the parameters selected for the analysis run. The counts change as you select values on each page that displays the Preliminary Counts section, allowing you to determine whether the preliminary set of patients potentially will yield meaningful outcome data before you perform the analysis run. The preliminary counts section displays more information as you proceed through the wizard. For each cohort, they can include potential patients with the index term, the outcome term, and incidence (both the index term and the outcome term).

	The preliminary counts do not use the specified analysis period or any temporal parameters of the run. Therefore, a patient counted in preliminary counts in the run wizard might not be in counts in the actual analysis results.
Select outcome variable	Outcome event variable. Available variables are determined by the configuration you are using.
If selecting more than one value, provide name for custom term	Name for the custom term defined as the index event. If you select multiple values in the list box, you must provide a name.
[List box for selecting values]	See Selecting values.
Remove terms if they are not in the source data	When selected, terms that are in the list of specified values and that are not in the source data are removed from the list when you continue to the next page of the wizard.
Exposure extension	(Available for a drug variable.) When selected, you can specify an exposure extension in days.

Step 3—Defining breakdowns for results

Previous step	One of the following: • Step 2c—Specifying temporal parameters. • Step 2—Defining an outcome event (Drug Utilization or Outcome Characterization).
Relevance	All types of descriptive analysis runs.
Description	In this step, you define breakdowns of analysis results. The results table contains one row for each breakdown. The breakdown value is in the SUBSET column of results.

To define breakdowns of analysis results:

1. On the Descriptive Analysis: Define Breakdowns for Results page, fill in the fields according to the following table, and click **Next**.
2. One of the following pages appears:
   - If you selected One or more co-occurrence windows, the Specify Parameters for Co-Occurrence page appears. Continue to Step 4—Specifying parameters for a co-occurrence window.
   - If you did not select **One or more co-occurrence windows** and you have the Access Advanced Run Options permission, the Descriptive Analysis: Advanced Options page appears. Continue to Step 5—Specifying advanced options.
   - If you did not select **One or more co-occurrence windows** and you do not have the Access Advanced Run Options permission, the Run Options page appears. Continue to Step 6—Specifying run options.

Field descriptions—Descriptive Analysis: Define Breakdowns for Results page

Field	Description
Break down the analysis results by the following	
<Demographic variables>	Demographic variables as determined by the configuration. For example, Gender is typically available as a breakdown. In the analysis results, one subset row appears for each value of the demographic variable that any patients in the initial cohort have. For example, when a patient group of only males is used as the initial cohort and the results are broken down by gender, there is only an ALL subset and a subset for males. In the SUBSET column of analysis results, the breakdown appears as ***<column name>*:** followed by the variable value for that breakdown, such as **GENDER: F**. If the value is not in the source data, the breakdown value is **null**.

Custom Age Group (of computed age)	(Appears if the configuration is set up to support custom age groups.) The computed age group is based on a computed age and on age groups that you define. In analysis results, one subset row appears for each computed age group that you define. Patients are in the age group if their computed age is within that range of the age group. In the SUBSET column of analysis results, the breakdown appears as **AGEGRP (Custom):** followed by the computed age group. Notes: • If the age group cannot be computed, the breakdown appears as **AGEGRP (Custom): Unknown**. • If the computed age is a negative number, the breakdown appears as **AGEGRP (Custom): Less than 0**. For more information, see Working with age groups.
Time category	Break down results for the following time categories: • **Year**—Create a breakdown for each calendar year, starting in January, within the analysis period defined for the analysis. In the analysis results, the subset names are **YEAR: *<year>***. • **Quarter**—Create a breakdown for each quarter, starting in January, within the analysis period defined for the analysis. In the analysis results, the subset names are **QUARTER: *<year and quarter>***. • **Month**—Create a breakdown for each month within the analysis period defined for the analysis. In the analysis results, the subset names are **MONTH: *<year and month>***.
One or more co-occurrence windows	When selected, the Specify Parameters for Co-Occurrence Window page appears later in the wizard, allowing you to identify co-occurring values during a specified time period. For more information, see Step 4—Specifying parameters for a co-occurrence window.

Step 4—Specifying parameters for a co-occurrence window

Previous step	Step 3—Defining breakdowns for results.
Relevance	Any type of descriptive analysis, as long as you selected **One or more co-occurrence windows** when defining breakdowns for results.
Description	In this step, you decide whether to identify co-occurring values, such as co-morbidities, medical history, and concomitant medications, during a specified time period, which is the co-occurrence window. You can define multiple co-occurrence windows. A co-occurrence window is a type of breakdown of results, resulting in subset rows for co-occurrence windows in the analysis results. For a co-occurrence value to appear as a row in results, at least one patient must have the co-occurring value. To define a co-occurrence window, you perform the following tasks: • Define a time period relative to a reference date. • Select a variable to be used for the co-occurrence window. • Either explicitly select co-occurring values to look for, or request breakdowns by the most frequent values within the co-occurrence window. If the co-occurrence window is based on a: • Diagnosis or procedure variable—Diagnoses or procedures that occurred within the window are considered to be co-occurring. • Drug variable—Drugs with any part of a drug era during the window are considered to be co-occurring.

To specify parameters for a co-occurrence window:

1. On the Specify Parameters for Co-occurrence Window page, fill in the fields according to the following table, and click **Next**.

2. To delete the co-occurrence window, click **Delete**, located at the bottom of the page, near the Back and Next buttons.

   A confirmation message appears. If you confirm the deletion, the current window is deleted. If you delete all co-occurrence windows, the **One or more co-occurrence windows** checkbox on the Define Breakdowns for Results page is deselected.

3. To add another co-occurrence window, click **Add**, located at the bottom of the page, near the Back and Next buttons.

   The window is added immediately after the current window. The current window is saved and a new, blank Specify Parameters for Co-occurrence Window page appears. You can define an unlimited number of co-occurrence windows.

4. Click **Next**.

One of the following pages appears:

- If you added another co-occurrence window, the Specify Parameters for Co-occurrence Window page appears. Fill in the fields according to the following table.
- If you did not add a co-occurrence window:
  - If you have the Access Advanced Run Options permission, the Descriptive Analysis: Advanced Options page appears. Continue to Step 5—Specifying advanced options.
  - If you do not have the Access Advanced Run Options permission, the Run Options page appears. Continue to Step 6—Specifying run options.

Field descriptions—Specify Parameters for Co-occurrence Window page

Field	Description
Window name (prefix)	Name of the co-occurrence window. The values that you select later on the page determine the text that appears in the SUBSET column in the analysis results. If you: • Select values to include or exclude, the column displays the window name. • Select **Enumerate ___ most frequently occurring values**, the following text appears: ***<window-name>: <frequent value>***.
Window description	Description of the co-occurrence window.
Select window variable	Co-occurrence variable. Available variables are determined by the configuration you are using.
Define Co-occurrence Window	Window type. If necessary, you also specify a window start and stop. For more information, see Window types.
Enumerate ___ most frequently occurring values	When selected, co-occurring values are defined as the most frequent values within the co-occurrence window. Type a positive integer that is no greater than 50. For more information, see Most frequent values.
Select specific values Specify one or more inclusion values Specify zero or more	When selected, you can create a co-occurrence window to find the following information: • Patients who have any of up to 25 inclusion values during the co-occurrence window.

exclusion values

- Patients who do not have any of up 25 exclusion values during the co-occurrence window.
- Patients who have any of up to 25 inclusion values during the window and do not have any of up to 25 exclusion values during the window.

You specify the values of the co-occurrence variables.

The selected values are treated as a group for which there is a single breakdown. The name of the breakdown in the SUBSET column of results is ***<window-name>***.

Inclusion values

A patient is considered to have the co-occurrence breakdown if any of the inclusion values are in the co-occurrence window. For example, if you select DrugA, DrugB, and DrugC as co-occurring values and name the co-occurrence window ABC, the analysis results contain a single row for which SUBSET is ABC, representing patients who had any of DrugA, DrugB, or DrugC during the co-occurrence window.

Note: If you select the same value that you used to define the outcome event or index event, the value appears as a co-occurring value.

Exclusion values

You can also select exclusion values. If you select only exclusion values, a patient is considered to have the co-occurrence breakdown only if none of the exclusion values are experienced by the patient in the co-occurrence window. For example, if the exclusion values are DrugD, DrugE, and DrugC, the patient cannot have any of those drugs in the window.

Both inclusion and exclusion values

If you specify both inclusion values and exclusion values, a patient has the co-occurrence breakdown only if the following criteria are met:

- The patient has any of the inclusion values in the window.
- The patient does not have any of the exclusion values in the window.

For example, if you specify DrugA, DrugB, and Drug C as inclusion values and DrugD, DrugE, and DrugF as exclusion values, a patient has the co-occurrence breakdown if the patient has any of the following circumstances:

- DrugA during the co-occurrence window and no DrugD, DrugE, or DrugF during the co-occurrence window.
- DrugB during the co-occurrence window and no DrugD, DrugE, or DrugF during the co-occurrence window.
- DrugC during the co-occurrence window and no DrugD,

	DrugE, or DrugF during the co-occurrence window. You cannot specify the same value as both an included value and an excluded value. **Selecting values** You can specify no more than 25 values. For information about selecting values, see Selecting values.
Remove terms if they are not in the source data	When selected, terms that are in the list of specified values and that are not in the source data are removed from the list when you continue to the next page of the wizard.
Exposure extension: ___ days	(Available for a drug variable.) When selected, you can specify an exposure extension in days.
No occurrence of co-occurrence term for ____ days before window start	(Available for a Risk-Outcome analysis run if an index lead-in is specified.) When selected, you can specify a co-occurrence lead-in. The co-occurrence lead-in indicates that the co-occurrence window is immediately preceded by at least the specified number of days clear of the co-occurrence value and within the patient's enrollment period. If the co-occurrence lead-in is not met, a co-occurrence window is not created. The clear days occur within the patient's enrollment period, but they do not need to be within the analysis period specified for the run. For example, if the analysis period starts 100 days after the patient's enrollment start, the days can be within those 100 days before the start of the analysis period. If co-occurrence values are based on a drug variable, this option uses the specified exposure extension option when looking for days clear of the drug. For explicitly included values, none of the selected values can have occurred for the specified number of days before the co-occurrence window. For example, suppose that you specify DrugA, DrugB, and DrugC for inclusion and specify a co-occurrence lead-in of 10. The co-occurrence window must be preceded by at least 10 days in which there are no drug era days for DrugA, DrugB, or DrugC. For explicitly excluded values, none of the selected values can have occurred for the specified number of days before the co-occurrence window. For example, suppose that you specify DrugD, DrugE, and DrugF for exclusion and specify a co-occurrence lead-in of 10. The co-occurrence window must be preceded by at least 10 days in which there are no drug era days for DrugD, DrugE, or DrugF. If you specify DrugA, DrugB, and DrugC for inclusion and DrugD, DrugE, and DrugF for exclusion, and specify a co-occurrence lead-in of 10, the co-occurrence window must be preceded by at least

10 days in which there are no drug era days for DrugA, DrugB, DrugC, DrugD, DrugE, or DrugF. Then, for the patient to have the co-occurrence breakdown, the co-occurrence window must include drug era days for any of DrugA, DrugB, or DrugC and must not include drug era days for any of DrugD, DrugE, or DrugF.

For most frequent values, a patient will not count as having a co-occurrence value unless the co-occurrence lead-in is met. For example, suppose 100 patients have DrugA within what would be the co-occurrence window (if the co-occurrence lead-in was met), but only 10 of the patients meet the co-occurrence lead-in. For the purposes of ranking the most frequently occurring values, DrugA is attributed with having a frequency of 10.

Day ranges

Some co-occurrence window options are for a specified day range relative to a reference date as follows:

- A positive integer indicates days after the reference date.
- A negative integer indicates days before the reference date.
- 0 indicates the reference date itself.

You can indicate a one-day co-occurrence window by entering the same number as the start and end of the window.

Window start and end

When the analysis runs, the start of a co-occurrence window is the later of the following options:

- Patient's enrollment start
- Specified window start

The co-occurrence window can include days that are before the analysis period start and after the patient's enrollment start.

If the specified window start is later than either the patient enrollment end or the analysis period end, a co-occurrence window is not created.

When the analysis runs, the end of a co-occurrence window is the earliest of the following options:

- Enrollment end
- Analysis period end
- Specified window end

Restrictions

Several restrictions ensure that, for a Risk-Outcome analysis, you cannot define a window that would start before a patient's enrollment start.

For a Risk-Outcome analysis, the values you enter must meet the

following criteria:

- If the co-occurrence window is based on the entire index exposure—The co-occurrence lead-in cannot be greater than the index lead-in.
- If the co-occurrence window is based on fixed start/stop relative to outcome date—The co-occurrence lead-in minus the co-occurrence window start cannot be greater than the greater of the index lead-in and the outcome lead-in.
- If the co-occurrence window is based on any other option—The co-occurrence lead-in minus the co-occurrence window start cannot be greater than the index lead-in.

If you do not specify an index lead-in, outcome lead-in, or co-occurrence lead-in, the value of each is considered to be 0. For example, if you do not specify an index lead-in and the co-occurrence window is based on index date, you cannot enter a window start of less than 0.

Step 5—Specifying advanced options

Previous step	One of the following: • Step 3—Defining breakdowns for results. • Step 4—Specify parameters for a co-occurrence window.
Relevance	All types of descriptive analysis runs, but only if you have the Access Advanced Run Options permission.

To specify advanced options:

1. On the Descriptive Analysis: Advanced Options page, fill in the fields according to the following table, and click **Next**.

   The Run Options page appears.

2. Continue to Step 6—Specifying run options.

Field descriptions—Descriptive Analysis: Advanced Options page

Field	Description
Maximum memory available to metrics-generation engine	Maximum amount of memory to devote to the engine that generates metrics. The initial value is the lesser of following values: • The estimate of the memory required for the run. The estimate is fixed at 8.0 GB. • The maximum allowed value as determined during setup.

Step 6—Specifying run options

Previous step	One of the following: • Step 3—Defining breakdowns for results. • Step 4—Specifying parameters for a co-occurrence window. • Step 5—Specifying advanced options.
Relevance	All types of descriptive analysis runs.

To specify run options:

1. On the Run Options page, fill in the fields according to the following table, and click **Next**.

   The Name Analysis Run page appears.

2. Continue to Step 7—Naming an analysis run.

Field descriptions—Run Options page

Field	Description
Run as soon as possible Do not run until ___ at ___	When to execute the descriptive analysis run. Select **Run as soon as possible** to execute the run immediately, or specify the date and time to create the run.
Email me when complete	When selected, an email message is sent to the email address you provide after the run either completes successfully or fails. Separate multiple email addresses with commas.
Save intermediate data files	(Available if you have the Save Intermediate Files permission.) When selected, supplemental data files that are produced when a run is created are saved for viewing and downloading.

Step 7—Naming an analysis run

Previous step	Step 6—Specifying run options.
Relevance	All types of descriptive analysis runs.

To name the analysis run:

1. On the Name Analysis Run page, fill in the fields according to the following table, and click **Next**.

   The Confirm Analysis Run Parameters page appears.

2. Review the selections for the run, and click **Submit Run**.

   A confirmation page appears, indicating that the run has been submitted. You can check the status of the run in the Status column.

   For information about results, see Viewing the results of a descriptive analysis run.

Field descriptions—Name Analysis Run page

Field	Description
Run name Description	Name and description of the run.
Add to	**Existing project**—The run is added to the project selected in the drop-down list. **New project named**—A new project is created with the name you provide, and the run is added to it.

Additional information for creating a descriptive analysis run

Window types

This information relates to choosing a window type in Step 4—Specifying parameters for a co-occurrence window.

Risk-Outcome analysis—Available window types

Option	Description
Fixed start/stop relative to index date (default)	The co-occurrence window starts and ends on specified days relative to the index date. If no index date is determined, a co-occurrence window is not created.

Entire index exposure	(Available if the index is based on the start of the first occurrence of a drug or the start of the last occurrence of a drug.)
	• If the index is the start of first occurrence:
	  • The window starts on the index date.
	  • The window ends at the end of the drug era that encompasses that index date.
	• If index is start of last occurrence:
	  • The window starts on the index date. The system determines the prescription that determined the index date.
	  • The window ends at the end of the drug era that encompasses that prescription. Drug eras are used to determine the end of the index exposure period and thus the end of the co-occurrence window, even though drug eras are not used to determine the index date itself.
	If no index date is determined, a co-occurrence window is not created.

Fixed start/stop relative to outcome date	The co-occurrence window starts and ends on specified days relative to the date of the first occurrence of the outcome event within follow-up. If the outcome event is a drug, the outcome date is the first day of the first drug era that starts within follow-up. If no outcome date is determined, a co-occurrence window is not created.
Entire follow-up	The co-occurrence window is the same as the period of follow-up. You select follow-up options when you specify temporal parameters.

Fixed start/stop relative to start of follow-up	The co-occurrence window starts and ends on a specified day relative to the start of follow-up. You select follow-up options when you specify temporal parameters. Follow-up end is ignored. If no index date is determined, a co-occurrence window is not created.
Fixed start/stop relative to end of follow-up	The co-occurrence window starts and ends on specified days relative to the end of follow-up. You select follow-up options when you specify temporal parameters. Follow-up start is ignored. If no index date is determined, a co-occurrence window is not created.

Drug Utilization analysis—Available window types

Option	Description
Fixed start/stop relative to index date (default)	The co-occurrence window starts and ends on specified days relative to the index date. If no index date is determined, a co-occurrence window is not created.
Entire index exposure	The window starts on the index date and ends at the end of the drug era that encompasses that index date. If no index date is determined, a co-occurrence window is not created.

Outcome Characterization analysis—Available window types

Option	Description
Fixed start/stop relative to index date	The co-occurrence window starts and ends on specified days relative to the index date. If no index date is determined, a co-occurrence window is not created.

Most frequent values

This information relates to Step 4—Specifying parameters for a co-occurrence window.

When you create an analysis run and select **Enumerate ___ most frequently occurring values**, co-occurring values are defined as the most frequent values within the co-occurrence window. The name of the breakdown appears in the SUBSET column of results with the following name: ***<window-name>: <frequent value>***.

If the co-occurrence is based on the same variable as the outcome event, the search for frequent values excludes values that were specified for the outcome event. This exclusion occurs regardless of the co-occurrence window type. For example, if the outcome event variable and the co-occurrence variable are Generic Name, and **Aspirin** is the outcome event, the search for most frequent values excludes **Aspirin**.

If the co-occurrence is based on the same variable as the index event, the search for frequent values excludes values that were specified for the index event. This exclusion occurs regardless of the co-occurrence window type.

To determine frequency, the application does the following tasks:

1. Counts the number of patients who have each value at least once during the co-occurrence window and for which the co-occurrence lead-in criteria is met. A patient is counted for a value only once.
2. Determines the values that have the highest patient counts, up to the limit of most frequent values that you specified.

When the limit of the most frequent value is reached, if there are multiple values with that same frequency, they are all included as most frequent values.

For example, suppose that the frequencies of drugs are as follows.

Drug name	Frequency (number of patients with the drug in the window)
DrugA	200
DrugB	200
DrugC	150
DrugD	150
DrugE	100
DrugF	100
DrugG	100

The following table shows the drugs that are identified as most frequent values depending on the specified number of most frequent terms. The most frequent values shown in the second column of the table below appear as subsets in analysis results.

Specified number of most frequent values	Most frequent values returned
1	DrugA DrugB
2	DrugA DrugB
3	DrugA DrugB DrugC DrugD
4	DrugA DrugB DrugC DrugD
5	DrugA DrugB DrugC DrugD DrugE DrugF DrugG

Now suppose that DrugD was specified as the index event in the analysis. DrugD would be ignored by the frequency counts, so the counts would be as follows.

Drug name	Frequency (number of patients with the drug in the window)
DrugA	200
DrugB	200
DrugC	150
DrugE	100

DrugF	100
DrugG	100

The following table shows the drugs that are identified as most frequent values depending on the user-specified number of most frequent terms.

Specified number of most frequent values	Most frequent values returned
1	DrugA DrugB
2	DrugA DrugB
3	DrugA DrugB DrugC
4 or 5	DrugA DrugB DrugC DrugE DrugF DrugG

Results of a descriptive analysis run

Viewing the results of a descriptive analysis run

Prerequisites

- You must have the View Analysis Results permission.

Overview

After a run completes, you can view the results in a table, which includes information such as:

- Counts of patients and days at risk.
- Statistics, such as incidence rate, that are computed from patient counts.

If the run was specified to include breakdowns of results, the table also includes separate rows for the individual breakdowns. For descriptions of the fields in the results table, see:

- Results of a Risk-Outcome analysis.
- Results of an Outcome Characterization analysis.
- Results of a Drug Utilization analysis.

To view analysis results:

1. Select the **Descriptive Analysis** tab.
2. To filter the analysis results that appear, select options from the **Project** and **Configuration** drop-down lists.
3. Navigate to the row containing the run for which to view results.

   Note: The run must have completed successfully. To verify that the run has completed, check the **Status** column.

4. Select the row menu (), and select **View Results**.
5. A page containing the results of the run appears. The page name varies, depending on the type of run.

   The results table contains the following rows:

   - One row for which ALL appears in the SUBSET column.
   - One row for each breakdown, if any, that was specified for the run and that has at least one patient counted in any column. The SUBSET column identifies the breakdown.

   Several columns include patient counts with links that allow you view details about the patients. For more information, see About drilldown.

6. To view details about the run, click **Show Notes**. Hide the notes by clicking **Hide Notes**.

Results of a Risk-Outcome analysis

Point to a column to view a tooltip.

Field descriptions—Results table for a Risk-Outcome analysis

Column	Description for the ALL subset	Description for other subsets, if different from the ALL subset
% N(Index)	**100.0** appears.	The following calculation is used: [N(Index) for subset / N(Index) for ALL subset] x 100
% N(Index) Overall	The following calculation is used: [N(Index) / N(Initial)] x 100	The following calculation is used: [N(Index) for subset / N(Initial) for ALL subset] x 100
% N(Initial)	**100.0** appears.	The following calculation is used: [N(Initial) for subset / N(Initial) for ALL subset] x 100
% N(Prvl)	**100.0** appears.	The following calculation is used: [N(Prvl) for subset / N(Prvl) for ALL subset] x 100
% N(Prvl) Overall	The following calculation is used: [N(Prvl) / N(Initial)] x 100	The following calculation is used: [N(Prvl) for subset / N(Initial) for ALL subset] x 100
AVGDAYS	(Available when the outcome event is a drug variable.) For patients counted in N(Initial), average days supplied for prescriptions that are counted in TOTRX. The following calculation is used: TOTSUPDAYS / TOTRX	The following calculation is used: TOTSUPDAYS for the subset / TOTRX for the subset
AVGDUR	(Available when the outcome event is a drug variable.) For patients counted in N(Initial), the average number of days in drug eras that are counted in TOTERA. The following calculation is used: TOTERADAYS / TOTERA	The following calculation is used: TOTERADAYS for the subset / TOTERA for the subset

CoWin #	Always empty.	Number of the co-occurrence window, if co-occurrence windows were specified in the run and appear in results. The number is automatically assigned and is visible in the wizard when you create the analysis.
DIM	**2** appears.	
I_025	Lower 2.5% confidence limit for the incidence proportion (INCIDENCE value).	
I_05	Lower 5% confidence limit for the incidence proportion (INCIDENCE value).	
I_95	Upper 95% confidence limit for the incidence proportion (INCIDENCE value).	
I_975	Upper 97.5% confidence limit for the incidence proportion (INCIDENCE value).	
ID	Automatically assigned unique row number.	
INCIDENCE	Proportion of patients who have an index date and who also had the outcome event during follow-up. The following calculation is used: N(Incidence) / N(Index)	The following calculation is used: N(Incidence) for subset / N(Index) for subset
IR	Incidence rate, which is the rate of occurrence of the outcome event per year. The following calculation is used: N(Incidence) / Person Years	The following calculation is used: N(Incidence) for subset / Person Years for subset
IR_025	Lower 2.5 percent confidence limit for the incidence rate.	
IR_05	Lower 5 percent confidence limit for the incidence rate.	
IR_95	Upper 95 percent confidence limit for the incidence rate.	

IR_975	Upper 97.5 percent confidence limit for the incidence rate.	
N(Incidence)	**If the outcome event is a diagnosis or procedure** Number of patients who have that diagnosis or procedure within follow-up. **If the outcome event is a drug** Number of patients who have the start of a drug era for that drug within follow-up. If a drug era starts before follow-up and continues into follow-up, the patient is not counted as having the outcome event during follow-up unless the patient has a separate, subsequent drug era that starts within follow-up. This rule applies regardless of whether an exposure extension was specified for the outcome event. For example, suppose that the follow-up start is October 1, 2005 and a patient has only two prescriptions for the drug: • September 28, 2005, 10-day supply. • October 3, 2005, 10-day supply. These prescriptions make up one drug era that is from September 28, 2005 through October 12, 2005. The patient is not counted in N(Incidence) because the patient does not have the start of a drug era during follow-up. If a separate era started during follow-up, the patient would be counted in N(Incidence).	Number of patients who are counted in N(Incidence) for the ALL subset and who meet the following criteria: **For breakdown by year/quarter/month** Patients whose outcome event during follow-up is within the year/quarter/month. Only the start of a drug era is considered. **For co-occurrence breakdown** Patients have the co-occurring value in the co-occurrence window. **For demographic breakdown** Patients have the demographic characteristic.
N(Index)	Number of patients with an index date.	Number of patients who are counted in N(Index) for the ALL subset and who meet the following criteria: **For breakdown by year/quarter/month** Patients have any part of follow-up within the year/quarter/month. This

		count uses follow-up days, which might start later than the index date. For example, suppose the time category is Month. If the index is on March 1, 2005 and follow-up start is on June 1, 2005, the patient is not in N(Index) for March, April, or May of 2005 because follow-up does not start until June. Follow-up end depends on how follow-up was defined and whether the outcome event occurs during follow-up. **For co-occurrence breakdown** The co-occurrence window is any time except start/stop relative to the outcome, and the patient has the co-occurring value in the co-occurrence window. This column is empty if the co-occurrence window is based on start/stop relative to outcome. **For demographic breakdown** Patients have the demographic characteristic.
N(Initial)	Number of patients in the initial cohort who have any part of their enrollment period during the analysis period.	Number of patients who are counted in N(Initial) for the ALL subset and meet the following criteria: **For breakdown by year/quarter/month** Patients have any enrollment days within the year/quarter/month. **For co-occurrence breakdown** This column is always empty. **For demographic breakdowns** Patients have the demographic characteristic.
N(Prvl)	**If the outcome event is a diagnosis or a procedure** Number of patients who have the diagnosis or procedure at any time within both their enrollment period and the analysis period. **If the outcome event is a drug**	Number of patients who are counted in N(Prvl) for the ALL subset and meet the following criteria: **For breakdown by year/quarter/month** Any of the values (diagnoses, procedures, or drug era days) for the outcome event are within the

	Number of patients with any part of a drug era for that drug within both their enrollment period and the analysis period. The drug era must start within the patient's enrollment period but does not necessarily need to start within the analysis period specified for the run.	year/quarter/month. **For demographic breakdown** Patients have the demographic characteristic. For co-occurrence breakdowns, this column is always empty.
Person Years	Total number of person years in follow-up. The following calculation is used: person-days / 365.24 Follow-up start and end depend on how follow-up was defined for the analysis. Person Years are from follow-up start up to and including the last day of follow-up. For patients counted in N(Incidence), the last day of follow-up is the date of the outcome event. You select follow-up options when you specify temporal parameters.	Total number of person years in follow-up that are counted for the ALL subset and that meet the following criteria: **For breakdown by year/quarter/month** Person-days are within the year/quarter/month. **For co-occurrence breakdown** Person-days are for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** Person-days are for patients who have the demographic characteristic.
PRVL	Period prevalence across the analysis period. The following calculation is used: N(Prvl) / N(Initial)	The following calculation is used: N(Prvl) for subset / N(Initial) for subset
SUBSET	**ALL** appears. If the breakdown is for a co-occurrence window for explicitly included or excluded values, you can point to the subset to view a tooltip showing the values. A subset row appears if the patient count for the subset is at least 1, even if the only column with at least one patient is N(Initial).	Name of the subset. One subset appears for each breakdown.
TOTERA	(Available when the outcome event is a drug variable.) For patients counted in N(Initial), the total number of drug eras that meet the following criteria:	Total number of drug eras that are counted in TOTERA for the ALL subset and that meet the following criteria: **For breakdown by**

	• Drug eras are for the drug specified as the outcome event. • Drug eras have any days within both the patient's enrollment period and the analysis period. This column can contain a non-zero value even if the TOTRX column is 0 because no prescriptions started during the enrollment period and analysis period.	**year/quarter/month** Drug eras have any days within the year/quarter/month. **For co-occurrence breakdown** Drug eras are for patients who have the co-occurring value in the co-occurrence window. This count is not of the co-occurring value. This count is for the drug specified as the outcome event, for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** Drug eras are for patients who have the demographic characteristic.
TOTERADAYS	(Available when the outcome event is a drug variable.) For patients counted in N(Initial), the sum of drug era days for drug eras that are counted in TOTERA. A drug era day is counted only once.	Sum of drug era days that are counted in TOTERADAYS for the ALL subset and that meet the following criteria: **For breakdown by year/quarter/month** Drug era days are within the year/quarter/month. A drug era day is counted only once. **For co-occurrence breakdown** Drug era days are for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** Drug era days are for patients who have the demographic characteristic.
TOTRX	(Available when the outcome event is a drug variable.) For patients counted in N(Initial), the total number of individual prescriptions that meet the following criteria: • Prescriptions are for the drug specified as the outcome event. • Prescriptions start within both the patient's enrollment period and the analysis	Total number of prescriptions that are counted in TOTRX for the ALL subset and that meet the following criteria: **For breakdown by year/quarter/month** Prescriptions start within the year/quarter/month. **For co-occurrence breakdown** Prescriptions are for patients who have the co-occurring value in the co-occurrence window. This count is

	period. (Prescriptions do not need to end within the enrollment period or analysis period.) If a patient has a single prescription that starts before the patient's analysis period and continues into it, 0 appears.	not of the co-occurring value. This count is for the drug specified as the outcome event, for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** Prescriptions are for patients who have the demographic characteristic.
TOTSUPDAYS	(Available when the outcome event is a drug variable.) For patients counted in N(Initial), the sum of days supplied for the prescription counted in TOTRX. All days supplied for a prescription are counted, even if the prescription continues past the end of the patient's enrollment period or the analysis period and even if the prescription overlaps with another prescription.	Sum of days supplied for prescriptions that are counted in TOTSUPDAYS for the ALL subset and that meet the following criteria: **For breakdown by year/quarter/month:** The prescriptions start within the year/quarter/month. All days supplied are counted for prescriptions that start within the time category, even if the prescriptions continue past the end of the time category or the patient’s enrollment period. **For co-occurrence breakdown** The prescriptions are for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** The prescriptions are for patients who have the demographic characteristic.

Results of an Outcome Characterization analysis

Point to a column to view a tooltip.

Field descriptions—Results table for an Outcome Characterization analysis

Column	Description for the ALL subset	Description for other subsets
% N(Initial)	**100.0** appears.	The following calculation is used: [N(Initial) for subset / N(Initial) for ALL subset] x 100
% N(Prvl)	**100.0** appears.	The following calculation is used: [N(Prvl) for subset / N(Prvl) for ALL subset] x 100
% N(Prvl) Overall	The following calculation is used: [N(Prvl) / N(Initial)] x 100	The following calculation is used: [N(Prvl) for subset / N(Initial) for ALL subset] x 100
CoWin #	Always empty.	Number of the co-occurrence window, if co-occurrence windows were specified in the run and appear in results. The number is automatically assigned and is visible in the wizard when you create the analysis.
DIM	**1** appears.	**1** appears.
ID	Automatically assigned unique row number.	Automatically assigned unique row number.
N(IncidentOutcome)	Always empty.	(A value appears only for breakdowns by year/quarter/month.) Count of patients with the outcome event in the year/quarter/month and within enrollment, but not in the prior year/quarter/month. Patients can be counted in this column only if they were enrolled for the entire prior year/quarter/month. Always empty for the first year/quarter/month.
N(Initial)	Number of patients in the initial cohort who have any part	Number of patients who are counted in N(Initial) for the ALL subset and meet the

	of their enrollment period during the analysis period.	following criteria: **For breakdown by year/quarter/month** Patients have any enrollment days within the year/quarter/month. **For co-occurrence breakdown** This column is always empty. **For demographic breakdowns** Patients have the demographic characteristic. **Note:** This count is not necessarily the same as the count of patients in the initial cohort.
N(Prvl)	Number of patients who have the diagnosis or procedure specified as the outcome event at any time within both their enrollment period and the analysis period.	Number of patients who are counted in N(Prvl) for the ALL subset and meet the following criteria: **For breakdown by year/quarter/month** Any of the diagnoses or procedures specified as the outcome event are within the year/quarter/month. **For co-occurrence breakdown** Patients have the co-occurring value in the co-occurrence window. **For demographic breakdown** Patients have the demographic characteristic.
PRVL	Period prevalence across the analysis period. The following calculation is used: N(Prvl) / N(Initial)	The following calculation is used: N(Prvl) for subset / N(Initial) for subset
SUBSET	**ALL** appears.	Name of the subset. One subset appears for each breakdown. **Note:** If the breakdown is for a co-occurrence window for explicitly included or excluded values, you can point to the subset to view a tooltip showing the values. A subset row appears if the patient count for the subset is at least 1, even if the only column with at least one patient is N(Initial).
TOTEVENT	For patients counted in N(Initial), the total number of	Total number of diagnoses or procedures that are counted in TOTEVENT for the ALL subset and meet the following criteria:

	diagnoses or procedures that meet the following criteria: • The diagnoses or procedures are specified as the outcome event. • The diagnoses or procedures are within both the patient's enrollment period and the analysis period. A diagnosis or procedure that occurs multiple times on the same day is counted multiple times in this column.	**For breakdown by year/quarter/month** The diagnoses or procedures occurred within the year/quarter/month. A diagnosis or procedure that occurs multiple times on the same day is counted multiple times in this column. **For co-occurrence breakdown** The diagnoses or procedures are for patients who have the co-occurring value in the co-occurrence window. This count is not of the co-occurring value. This count is of the diagnosis or procedure that was specified as the outcome event, for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** The diagnoses or procedures are for patients who have the demographic characteristic.

Results of a Drug Utilization analysis

Point to a column to view a tooltip.

Field descriptions—Results table for a Drug Utilization analysis

Column	Description for the ALL subset	Description for other subsets
% N(Initial)	**100.0** appears.	The following calculation is used: [N(Initial) for subset / N(Initial) for ALL subset] x 100
% N(Prvl)	**100.0** appears.	The following calculation is used: [N(Prvl) for subset / N(Prvl) for ALL subset] x 100
% N(Prvl) Overall	The following calculation is used: [N(Prvl) / N(Initial)] x 100	The following calculation is used: [N(Prvl) for subset / N(Initial) for ALL subset] x 100
AVGDAYS	For patients counted in N(Initial), average days supplied for prescriptions that are counted in TOTRX. The following calculation is used: TOTSUPDAYS / TOTRX	The following calculation is used: TOTSUPDAYS for the subset / TOTRX for the subset
AVGDUR	For patients counted in N(Initial), average number of days in drug eras that are counted in TOTERA. The following calculation is used: TOTERADAYS / TOTERA	The following calculation is used: TOTERADAY for the subset / TOTERA for the subset
CoWin #	Always empty.	Number of the co-occurrence window, if co-occurrence windows were specified in the run and appear in results. The number is automatically assigned and is visible in the wizard when you create the analysis.

DIM	**1** appears.	**1** appears.
ID	Automatically assigned unique row number.	Automatically assigned unique row number.
N(IncidentUse)	Always empty.	(A value appears only for breakdowns by year/quarter/month.) Number of patients who have a drug era for the outcome event that started in the year/quarter/month and within enrollment, but who have no part of a drug era for the drug in the prior year/quarter/month. Patients can be counted in this column only if they were enrolled for the entire prior year/quarter/month. Always empty for the first year/quarter/month.
N(Initial)	Number of patients in the initial cohort who have any part of their enrollment period during the analysis period.	Number of patients who are counted in N(Initial) for the ALL subset and meet the following criteria: **For breakdown by year/quarter/month** Patients have any enrollment days within the year/quarter/month. **For co-occurrence breakdown** This column is always empty. **For demographic breakdowns** Patients have the demographic characteristic. **Note:** This count is not necessarily the same as the count of patients in the initial cohort.
N(Prvl)	Number of patients with any part of a drug era for the drug specified as the outcome event within both their enrollment period and the analysis period. The drug era must start within the patient's enrollment period but does not necessarily need to start within the analysis period specified for the run.	Number of patients who are counted in N(Prvl) for the ALL subset and meet the following criteria: **For breakdown by year/quarter/month** Any drug era days for the drug specified as the outcome event are within the year/quarter/month. **For co-occurrence breakdown** Patients have the co-occurring value in the co-occurrence window. **For demographic breakdown** Patients have the demographic

		characteristic.
PRVL	Period prevalence across the analysis period. The following calculation is used: N(Prvl) / N(Initial)	The following calculation is used: N(Prvl) for subset / N(Initial) for subset
SUBSET	**ALL** appears.	Name of the subset. One subset appears for each breakdown. **Note:** If the breakdown is for a co-occurrence window for explicitly included or excluded values, you can point to the subset to view a tooltip showing the values. A subset row appears if the patient count for the subset is at least 1, even if the only column with at least one patient is N(Initial).
TOTERA	For patients counted in N(Initial), total number of drug eras that meet the following criteria: • Drug eras are for the drug specified as the outcome event. • Drug eras have any days within both the patient's enrollment period and the analysis period. This column could contain a non-zero value even if the TOTRX column is 0 because no prescriptions started during the enrollment period and analysis period.	Total number of drug eras that are counted in TOTERA for the ALL subset and meet the following criteria: **For breakdown by year/quarter/month** Drug eras have any days within the year/quarter/month. **For co-occurrence breakdown** Drug eras are for patients who have the co-occurring value in the co-occurrence window. This count is not of the co-occurring values. Instead, this count is for the drug specified as the outcome event, for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** Drug eras are for patients who have the demographic characteristic.
TOTERADAYS	For patients counted in N(Initial), the sum of drug era days for drug eras that are counted in TOTERA. A drug era day is counted only once.	Sum of drug era days that are counted in TOTERADAYS for the ALL subset and that meet the following criteria: **For breakdown by year/quarter/month** Drug era days are within the year/quarter/month. A drug era day is

		counted only once. **For co-occurrence breakdown** Drug era days are for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** Drug era days are for patients who have the demographic characteristic.
TOTRX	For patients counted in N(Initial), the total number of individual prescriptions that meet the following criteria: • Prescriptions are for the drug specified as the outcome event. • Prescriptions start within both the patient's enrollment period and the analysis period. (Prescriptions do not need to end within the enrollment period or analysis period.) If a patient has a single prescription that starts before the patient's analysis period and continues into it, 0 appears.	Total number of prescriptions that are counted in TOTRX for the ALL subset and meet the following criteria: **For breakdown by year/quarter/month** Prescriptions start within the year/quarter/month. **For co-occurrence breakdown** Prescriptions are for patients who have the co-occurring value in the co-occurrence window. This count is not of the co-occurring value. This count is for the drug specified as the outcome event, for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** Prescriptions are for patients who have the demographic characteristic.
TOTSUPDAYS	For patients counted in N(Initial), the sum of days supplied for the prescription counted in TOTRX. All days supplied for a prescription are counted, even if the prescription continues past the end of the patient's enrollment period or the analysis period and even if the	Sum of days supplied for prescriptions that are counted in TOTSUPDAYS for the ALL subset and meet the following criteria: **For breakdown by year/quarter/month** The prescriptions start within the year/quarter/month. (All days supplied are counted for prescriptions that start within the time category, even if the prescriptions continue past the end of the time category or the patient's enrollment period.) **For co-occurrence breakdown**

prescription overlaps with another prescription.	The prescriptions are for patients who have the co-occurring value in the co-occurrence window. **For demographic breakdown** The prescriptions are for patients who have the demographic characteristic.

Creating and running a definition file

For an existing analysis run, you can create a definition file that allows you to submit the analysis run automatically instead of through the Empirica Healthcare Analysis user interface. You can run the definition file as often as needed.

Prerequisites

- You must be a superuser.
- The analysis run must have completed.

To create and run a definition file:

1. Select one of the following tabs:
   - Descriptive Analysis
   - Evaluative Analysis
   - Analysis Setup
2. Select the row menu () for an analysis run, and select **Create Definition File**.

   The Run Definition File dialog box appears and contains the contents of the file.
3. Select the full text in the dialog box.
4. Paste the text into a text application, such as Notepad.
5. To modify the name of the analysis run, edit the **name** parameter.

   Note: Do not modify the values of any other parameters.
6. Save the document with an extension of **IN**, and close the document.
7. Copy the file to the correct directory:

   A. Open the webvdme.properties file, located in the following directory on the application server:

   $INSTALL_DIR/Healthcare/WEB-INF/classes

   B. Find the temp_dir parameter, and note the directory it references, such as:

   temp_dir=/u01/app/oracle/product/EHC/temp

   C. Navigate to the specified directory, and locate the **input** folder.

   For example, for the previous example, the directory is

   /u01/app/oracle/product/EHC/temp/input

   D. Place the IN file in the input directory.

The listener process checks the input directory every 2 to 30 seconds and executes any run definition files with the IN extension. While being processed, the file has an extension of PROC. The file extension changes to DONE when the execution of the run definition file is complete.

Publishing an analysis run

You can publish any run that you create. When you publish a run, the results are available to other users, including those who might not otherwise have permissions to view the results and those who might not have privileges to create runs and use the configuration. Publishing allows multiple users to review the results of data that is analyzed only a single time, reducing machine time overhead.

You can publish a run as soon as it is submitted.

1. Select the **Descriptive Analysis** or **Evaluative Analysis** tab.

2. Select the row menu () for a run, and select **Publish**.

Evaluative Analysis tab

Evaluative Analysis page

In the Empirica Healthcare Analysis application, you can perform several types of analysis, including evaluative analysis. For information about evaluative analysis runs, see Evaluative analysis runs.

On the Evaluative Analysis tab, you create evaluative analysis runs and view information about runs, including their results. The page lists the following analysis runs:

- Analysis runs that you created.
- Analysis runs that are published to you or your login group.

Note: Your view might include runs that are listed for configurations to which you have not been granted permission.

General activities

The following links and filters appear at the top of the page and affect the entire page.

- Create Evaluative Analysis
- Columns
- Print
- Download
- Select Rows

Row-specific activities

The following menu options are available from the row menu, located in the leftmost column of the table, and affect an individual row in the table. Your permissions determine the options that appear.

- View Results
- View Cohort Summary
- Rename
- Cancel (for runs that are in progress) or Delete (for completed runs)
- Re-run
- View Jobs for Run
- View Run Details
- Publish

Field descriptions—Evaluative Analysis page

Field	Description
Configuration	Name of the configuration used by the run.

Created	Date and time the run was created.
Created By	User who created the run.
Description	Description of the run.
End Date	Date and time the run ended. This column is empty until all of the run's component jobs are complete.
ID	Identifier that was assigned automatically to the analysis run when the run was submitted. Analysis run IDs are unique across the analysis tabs and are not reused when runs are deleted.
Name	Name of the run.
Project	Name of the project to which the run is assigned.
Run Type	Type of run.
Size	Size of the results table.
Start Date	Date and time the run started. For a scheduled run, the start date might be substantially later than the creation date of the run.
Status	This column is empty until all of the run's component jobs are complete. • **Completed**—All jobs completed successfully. • **Cancelled**—The run was cancelled. • **Error Occurred**—A job for the run failed, so the run failed. • **Completed with Warnings**—A warning appears because not enough information exists to produce meaningful results. To view more information about warnings, open the Jobs for Run page and look in the Status column.

Evaluative analysis runs

An evaluative analysis run compares the following:

- The incidence rate of an outcome event during a follow-up period that is based on a specified index event.
- The incidence rate of the same outcome event during a follow-up period that is based on a different specified index event.

For example, you might want to compare the incidence rate of a diagnosis within 60 days after the start of exposure to DrugA with the incidence rate of the same diagnosis occurring within 60 days after the start of exposure to DrugB.

The index events and outcome event can be drugs, diagnoses, or procedures, as long as the configuration supports them.

Screening process for primary and comparator cohorts

When you create an evaluative analysis run, you select a patient group as the initial primary cohort. For the initial comparator cohort, you can select a patient group or use all patients. The analysis then performs screening, resulting in the screened primary cohort and screened comparator cohort.

During screening, the analysis performs the following activities:

- For the primary cohort, any patients who had the comparator index term at any time during enrollment are removed from the patient group. The result is the screened primary cohort. Patients in the screened primary cohort do not necessarily have the primary index event.
- For the comparator cohort, any patients who had the primary index term at any time during enrollment are removed from the patient group, or, if you used all patients, from the set of all patients. The result is the screened comparator cohort. Patients in the screened comparator cohort do not necessarily have the comparator index event.

The screening process ignores the specified analysis period and any temporal parameters specified for the run.

In the results of an analysis run, the N(Initial) count shows the following information:

- For the primary cohort, patients in the screened primary cohort who had at least one day of enrollment during the specified analysis period.
- For the comparator cohort, patients in the screened comparator cohort who had at least one day of enrollment during the specified analysis period.

The analysis determines the following information:

- For the primary cohort, the patients in Group A who had an index date based on the primary index event, and which of those patients had the outcome event during follow-up.
- For the secondary cohort, the patients in Group B who had an index date based on the comparator index event, and which of those patients had the outcome event during follow-up.

Diagnoses and procedures

A diagnosis is recorded in the source data for a patient. For example, the diagnosis from insurance claims data might be a term from the ICD-9-CM.

A diagnosis or procedure is a time point consisting of a single day. Diagnoses and procedures are determined from the following information in the source data:

- Diagnosis term or procedure term.
- Date that represents the occurrence of the diagnosis or procedure.

If you select multiple diagnoses or multiple procedures as the index event or outcome event in a run, the analysis treats the values as if they are one term.

Prescriptions

A prescription is a time interval consisting of one or more days. Prescriptions are determined from the following information in the source data:

- Drug term.
- Date that represents the prescription start.

  For example, the prescription's dispensation date, or for mail-order prescriptions, the dispensation date plus two days.
- Days supplied.

The prescription start date is the first day of the prescription. For example, consider a prescription that starts on March 1, 2008 and is supplied for four days. The first date of the prescription is March 1, 2008, and the last day is March 4, 2008.

If you select multiple drugs as the index event or outcome event in a run, the analysis treats the values as if they are one term.

Drug exposures and drug eras

During analysis runs, drug eras are generated from prescriptions. A drug era is an interval of time during which a patient is considered to be continuously exposed to a specified drug or group of drugs. When analyzing drug information, an analysis run looks at either prescriptions or drug eras, depending on the context.

Note: Prescription data is from the source data. Drug eras are constructed from the source data during analysis runs but do not change the source data.

Prescriptions that start before the patient’s enrollment start are ignored even if they continue into the patient’s enrollment period. For example, consider the following situation:

- A ten-day prescription starts on June 1, 2006.
- The patient’s enrollment start is June 5, 2006.
- A three-day prescription starts on June 8, 2006.

The first drug era is from June 8, 2006 through June 10, 2006.

Drug eras are constructed from the source data as follows:

1. The end of each prescription is determined.

   For example, suppose that a prescription starts on June 1, 2006 and the days supplied is 10. The prescription ends on June 10, 2006.

2. If two or more prescriptions are contiguous or overlapping, they are combined into a single drug exposure.

3. If no exposure extension is specified, each resulting drug exposure is a drug era.

   If you specify an exposure extension (a number of days), the extension is applied as follows:

   A. The specified number of days is added to the end of each resulting drug exposure identified in the first step.

   B. If the number of days between the end of one prescription and the start of the next prescription is less than or equal to the specified number of days, the prescriptions are combined. If the number of days between the two prescriptions is greater than the specified number of days, the prescriptions are not combined.

The result is one or more drug eras.

In the following example, four drug eras are constructed from the patient's prescriptions in the source data:

The following example shows how prescriptions are combined if an exposure extension is specified. Two drug eras are constructed.

High-dimension propensity scores

For a logistic regression run, you have the option to use high-dimensional propensity scores (HDPS). This option enables the logistic regression to adjust for large numbers of factors in patient healthcare data. The determination of the empirical variables to be used in generating high-dimensional propensity scores is performed using the Pharmacoepidemiology Toolbox Software provided at http://www.hdpharmacoepi.org.

After the empirical variables are determined, the variables are used by the Empirica Healthcare Analysis application to generate propensity scores, which indicate the propensity of patients to have the primary index event (such as exposure to a particular drug) rather than the comparator index event (such as exposure to a comparator drug).

To generate propensity scores, a logistic regression is performed for each included propensity score model. The propensity scores can be used in either of the following ways:

- To create categorical variables (deciles of propensity scores) that are used as covariates in the final regression.
- To match cohorts and do a final regression using the matched cohorts.

Creating an evaluative analysis

Step 1—Creating an evaluative analysis run

Prerequisites

The application must have been set up so the Evaluative Analysis tab is visible, and you must have the View Analysis Results permission, the Create Analysis Run permission, or both permissions.

To create an evaluative analysis run:

1. Select the **Evaluative Analysis** tab.
2. Click **Create Evaluative Analysis**.

   The Create Evaluative Analysis page appears.
3. Fill in the fields according to the following table, and click **Next**.

   The Define Index Event for Primary Cohort page appears.
4. Continue to Step 2—Defining an index event for the primary and comparator cohorts.

Field descriptions—Create Evaluative Analysis page

Field	Description
Perform cohort study using the following model	• **Poisson regression**—Base the analysis run on a Poisson regression. • **Logistic regression**—Base the analysis run on a logistic regression. • **Use high-dimensional propensity scores**—When selected, you have the option to use high-dimensional propensity scores (HDPS) in the analysis run. • **Cox proportional hazards**—Base the analysis run on a Cox proportional hazards model.
Use ridge regression	When selected, Ridge regression is used to increase the likelihood of convergence. For an HDPS run, ridge regression is always used in the generation of propensity scores; this option affects only the final logistic regression for each propensity score model.
Configuration	Configuration to use for the analysis run. Select a configuration by choosing it from the drop-down list or by clicking **Browse**.
Specify	Patient groups used to define the primary cohort and comparator cohort.

primary cohort: Select patient group and Specify comparison cohort: Restrict to patient group	You are required to choose a patient group for the primary cohort. If you do not select **Restrict to patient group**, the comparator cohort is all patients. Select a patient group using one of the following options: • Choose the patient group from the drop-down list, which includes all patient groups based on the selected configuration. • Click **Browse**, and select a patient group from the dialog box. • To create a patient group, click **Browse Data** at the top of the page. The new patient group that you create is added to the patient group drop-down lists immediately. For more information, see Browsing data. During the analysis, some patients might be dropped from the analysis for the following reasons: • A patient's enrollment period does not overlap with the specified analysis period. For example, suppose that a patient is enrolled from January 1, 2006 to January 1, 2007, and the analysis period start is May 1, 2007. If the patient were included in the initial cohort, the patient would be dropped from the analysis and therefore would not be represented in the N(Initial) column of analysis results. • The screening process removes some patients. For more information, see Evaluative analysis runs.
Start of analysis period	(Available after you select a primary cohort.) First day of data to analyze for the run. The default value is the earliest enrollment start date for any patients who are in either the initial primary cohort or the initial comparator cohort. You can select a date that is later than the default but not earlier.
End of analysis period	(Available after you select a primary cohort.) Last day of data to analyze for the run. The default value is the latest enrollment end date for any patients who are in either the initial primary cohort or the initial comparator cohort. You can select a date that is earlier than the default but not later.

Step 2—Defining an index event for the primary and comparator cohorts

Previous step	Step 1—Creating an evaluative analysis run.
Description	In this step, you define index events for both the primary cohort and the comparator cohort. An index event is a list of values that the analysis run uses to establish an index date for each patient. Defining an index event consists of the following steps: 1. Selecting a drug, diagnosis, or procedure variable. 2. Selecting values for the variable.

To define index events for the primary and secondary cohorts:

1. On the Define Index Event for Primary Cohort page, fill in the fields according to the following table, and click **Next**.

   The Define Index Event for Comparator Cohort page appears.

2. Fill in the fields according to the following table, and click **Next**.

   The Define Outcome Event page appears.

3. Continue to Step 3—Defining an outcome event.

Option descriptions—Define Index Event for Primary Cohort page and Define Index Event for Comparator Cohort page

Option	Description
Show Preliminary Counts	View preliminary counts of patients to help determine if you are selecting appropriate values. The preliminary counts section dynamically displays counts of possibly eligible patients based on the parameters selected for the analysis run. The counts change as you select values on each page that displays the Preliminary Counts section, allowing you to determine whether the preliminary set of patients potentially will yield meaningful outcome data before you perform the analysis run. If enough information has been provided for the screening process to occur, the counts are for patients in the screened cohorts. Otherwise, they are for patients in the unscreened cohorts. For each cohort, they can include potential patients with the respective index term, the outcome term, and incidence (both the respective index term and the outcome term). The preliminary counts do not use the specified analysis period or any temporal parameters of the run. Therefore, a patient counted in preliminary counts in the run wizard might not be in counts in the actual analysis results.

Select index variable or Select comparator index variable	Index event variable for the primary cohort or comparator cohort, depending on the page. Available variables are determined by the configuration you are using. If a patient is in the screened primary cohort, the analysis establishes an index date for the patient using the primary index term and the primary index lead-in. In establishing an index date for patients in the primary cohort, the analysis ignores the comparator index event. If a patient is in the screened comparator cohort, the analysis establishes an index date for the patient using the comparator index term and the comparator index lead-in. In establishing an index date for patients in the comparator cohort, the analysis ignores the primary index event. You can base the primary index event and the comparator index event on different types of variables. For example, the primary index event can be a drug and the comparator index event can be a procedure.
Define index event as _____ of terms selected below	Index type. For a diagnosis or procedure variable, the following types are available: • start of first occurrence • start of last occurrence For a drug variable, the following types are available: • start of first occurrence • start of last occurrence • end of last occurrence If you are doing an HDPS run and you plan to select dimensions, select **start of first occurrence** so that you can specify index lead-ins. For information about each of these options, see Determination of the index date.
If selecting more than one value, provide name for custom term	Name for the custom term defined as the index event. If you select multiple values in the list box, you must provide a name.
[List box for selecting values]	See Selecting values.
Remove terms if they are not in the source data	When selected, terms that are in the list of specified values and not in the source data are removed from the list when you continue to the next page of the wizard.

Exposure extension: ___ days	(Available for a drug variable.) When selected, you can specify an exposure extension in days.

Step 3—Defining an outcome event

Previous step	Step 2—Defining an index event for the primary and comparator cohorts.
Description	In this step, you specify one or more values as the outcome event, which can be based on a drug, diagnosis, or procedure variable. The analysis attempts to find the occurrence of the outcome event during follow-up. • If the outcome event is based on a diagnosis or procedure variable, the analysis looks for the first occurrence of the outcome event during follow-up. • If the outcome event is based on a drug variable, the analysis looks for the first day of the first drug era that starts during follow-up.

To define an outcome event:

1. On the Define Outcome Event page, fill in the fields according to the following table, and click **Next**.

   One of the following pages appears:

   - If you created a logistic regression and selected **Use high-dimensional propensity scores** on the Create Evaluative Analysis page, the Select HDPS Dimensions page appears. Continue to Step 4—Selecting HDPS dimensions.
   - If you did not select **Use high-dimensional propensity scores** on the Create Evaluative Analysis page, the Specify Temporal Parameters page appears. Continue to Step 5—Specifying temporal parameters.

Option descriptions—Define Outcome Event page

Option	Description
Show Preliminary Counts	View preliminary counts of patients to help determine if you are selecting appropriate values. The preliminary counts section dynamically displays counts of possibly eligible patients based on the parameters selected for the analysis run. The counts change as you select values on each page that displays the Preliminary Counts section, allowing you to determine whether the preliminary set of patients potentially will yield meaningful outcome data before you perform the analysis run. If enough information has been provided for the screening process to occur, the counts are for patients in the screened cohorts. Otherwise, they are for patients in the unscreened cohorts. For each cohort, they can include potential patients with the respective index term, the outcome term, and incidence (both the respective index term and the outcome term).

	The preliminary counts do not use the specified analysis period or any temporal parameters of the run. Therefore, a patient counted in preliminary counts in the run wizard might not be in counts in the actual analysis results.
Select outcome variable	Outcome event variable. Available variables are determined by the configuration you are using.
If selecting more than one value, provide name for custom term	Name for the custom term defined as the outcome event. If you select multiple values in the list box, you must provide a name.
[List box for selecting values]	See Selecting values.
Remove terms if they are not in the source data	When selected, terms that are in the list of specified values and not in the source data are removed from the list when you continue to the next page of the wizard.
Exposure extension	(Available for a drug variable.) When selected, you can specify an exposure extension in days.

Step 4—Selecting HDPS dimensions

Previous step	Step 3—Defining an outcome event.
Relevance	Appears only if you selected **Use high-dimensional propensity scores** on the Create Evaluative Analysis page.
Description	In this step, you select one or more dimensions that are evaluated for their possible influence on the propensity of patients to experience the primary index event.

To select HDPS dimensions:

1. On the Select HDPS Dimensions page, fill in the fields according to the following table, and click **Next**.

   The Specify Temporal Parameters page appears.

2. Continue to Step 5—Specifying temporal parameters.

Field descriptions—Select HDPS Dimensions page

Field	Description
Dimension name	Dimension to be used in the computation of high-dimensional propensity scores.
Select terms to exclude for dimension	Terms to exclude from each dimension. The following terms are excluded automatically from dimensions: • Terms that are used for the primary index event, comparator index event, or outcome event. • Terms on whose primary path lies the higher-level term used for the primary index event, comparator index event, or outcome event.

Step 5—Specifying temporal parameters

Previous step	One of the following: • Step 3—Defining an outcome event. • Step 4—Selecting HDPS dimensions.
Description	In this step, you specify a primary index lead-in, a comparator index lead-in, and an outcome lead-in. You also specify a period of follow-up, also known as a risk window, which is always relative to the index date.

To specify temporal parameters:

1. On the Specify Temporal Parameters page, fill in the fields according to the following table, and click **Next**.

   The Define Covariates page appears.

2. Continue to Step 6—Defining covariates.

Option descriptions—Specify Temporal Parameters page

Option	Description
Show Preliminary Counts	View preliminary counts of patients to help determine if you are selecting appropriate values. The preliminary counts section dynamically displays counts of possibly eligible patients based on the parameters selected for the analysis run. The counts change as you select values on each page that displays the Preliminary Counts section, allowing you to determine whether the preliminary set of patients potentially will yield meaningful outcome data before you perform the analysis run. If enough information has been provided for the screening process to occur, the counts are for patients in the screened cohorts. Otherwise, they are for patients in the unscreened cohorts. For each cohort, they can include potential patients with the respective index term, the outcome term, and incidence (both the respective index term and the outcome term). The preliminary counts do not use the specified analysis period or any temporal parameters of the run. Therefore, a patient counted in preliminary counts in the run wizard might not be in counts in the actual analysis results.
No occurrence of primary index term *<index term>* for ___ days before index date	(Available when the index date is based on the first occurrence of a diagnosis, procedure, or drug.) When selected, you can specify an index lead-in, which is the number of days before the index date that must be clear of the value specified as the index event. If the index event is a drug, the days must be clear of any part of a drug era for the drug.

and No occurrence of comparator index term *<index term>* for ___ days before index date	The clear days: • Occur immediately before the index event and are counted backwards from the index event, treating the index date as Day 0. • Occur within the patient's enrollment period, but they do not need to be within the analysis period specified for the run. For example, if the analysis period starts 100 days after the patient's enrollment start, the days can be within those 100 days before the analysis period start. The *<index term>* value in the label is one of the following: • The single value you specified. If the value is too long to fit, you can point to the underlined value to see a tooltip. • The custom term you specified. Point to the underlined custom term to see its values. If you do not specify an index lead-in, there is no requirement for a clear period before the index date. If you do not specify both a primary index lead-in and a comparator index lead-in, you are not able to specify a history window lead-in. For an HDPS run for which you selected dimensions, you must specify both a primary index lead-in and a comparator index lead-in.
No occurrence of outcome term *<outcome term>* for ___ days before follow-up start	When selected, you can specify an outcome lead-in, which is the number of days before follow-up start that must be clear of the value specified as the outcome event. If the outcome event is a drug, the days must be clear of any part of a drug era for the drug. The clear days: • Occur immediately before the follow-up start and are counted backwards from the follow-up start, treating the index date as Day 0. These days are before follow-up start, which could start after the index date. • Occur within the patient's enrollment period, but they do not need to be within the analysis period specified for the run. For example, if the analysis period starts 100 days after the patient's enrollment start, the days can be within those 100 days before the analysis period start. The *<outcome term>* value in the label is one of the following: • The single value you specified. If the value is too long to fit, you can point to the underlined value to see a tooltip. • The custom term you specified. Point to the underlined custom term to see its values. If you do not specify an outcome lead-in, there is no requirement for a clear period before the follow-up start.
Specify start of follow-up (risk window) as	• **Date of index event**—When selected, follow-up starts on and includes the index date. • **Follow-up begins ___ days after index event**—When selected, follow-up starts on the specified number of days after the index

	date. The index date is day 0, so a value of 1 is the day after the index date. For example, when a patient's index date is June 1, 2006 and you specify a follow-up start of 5 days after the index event, follow-up starts on June 6, 2006. This option does not preclude the possibility of an outcome event occurring after the index and before the follow-up start.
Specify end of follow-up (risk window) as earliest of these options Follow-up ends for a patient when any of the following selected options occurs.	
after ___ days of follow-up	When selected, follow-up ends after the specified number of days. For example, if you specify follow-up start as the index date, with an index date of June 1, 2006, and you specify 5 for this value, the last day of follow-up is June 5, 2006. If you specify follow-up start as 5 days after the index event, the first day of follow-up is June 6, 2006. If you specify 6 for the **after ___ days of follow-up** value, the follow-up end is June 11, 2006.
end of index exposure period	(Available only if both the primary index event and the comparator index event are based on the start of the first occurrence of a drug or the start of the last occurrence of a drug.) When selected: • If the index date is the start of the first occurrence of a drug exposure, follow-up ends on the last day of the drug era. • If the index date is the start of the last occurrence of a drug exposure, follow-up ends on the last day of the drug era that encompasses the prescription that determines the index date. Drug eras are used in determining the end of index exposure period and thus the end of follow-up, even though they are not used to determine the index date itself. If the follow-up start is specified in a way that would make follow-up start after the end of the drug era for the candidate index exposure, the index candidate is rejected. For example, if the drug era for the candidate index is from May 1, 2005 to May 5, 2005, and the follow-up start is specified as 5, the index candidate is rejected. The index date is counted as day 0, so 5 days after index is May 6, 2005.
upon first occurrence of outcome event *<outcome event>*	(Always selected; cannot be unselected.) • If the outcome event is a diagnosis or procedure, follow-up ends on the date of the first occurrence of the outcome event during follow-up. • If the outcome event is a drug, follow-up ends on the first day of the first drug era that starts during follow-up. If the outcome event is based on a drug variable and a drug era starts before follow-up and continues into follow-up, the patient is not counted as having the outcome event during follow-up unless the patient has a separate, subsequent drug era that starts within

	follow-up, regardless of whether an exposure extension is used.
upon end of analysis period or enrollment	(Always selected; cannot be unselected.) Follow-up ends at the end of the patient's enrollment period or the end of the specified analysis period for the analysis run.

Step 6—Defining covariates

Previous step	Step 5—Specifying temporal parameters.
Description	In this step, you can define covariates for the analysis to use. In the results of the analysis run, the Cohort Summary table contains a row for each combination of covariates, with the SUBSET column indicating the combination. A row does not appear in the table for a logistic run that uses the HDPS option.

- On the Define Covariates page, fill in the fields according to the following table, and click **Next**.

  One of the following pages appears:

  - If you selected **One or more covariates defined as history windows**, the Specify Parameters for History Window page appears. Continue to Step 7—Specifying parameters for the history window.
  - If you did not select **One or more covariates defined as history windows:**
    - If the analysis run is based on a logistic regression and you selected **Use high-dimensional propensity scores** on the Create Evaluative Analysis page, the Specify HDPS Parameters page appears. Continue to Step 8—Specifying HDPS parameters.
    - If the analysis run is based on the Cox proportional hazards model, the Evaluative Analysis: Cox Options page appears. Continue to Step 9—Specifying Cox options.
    - If the analysis run is based on a Poisson regression or a non-HDPS logistic regression and you have the Access Advanced Run Options permission, the Evaluative Analysis: Advanced Options page appears. Continue to Step 10—Specifying advanced options.
    - If the analysis run is based on a Poisson regression or a non-HDPS logistic regression and you do not have the Access Advanced Run Options permission, the Run Options page appears. Continue to Step 11—Specifying run options.

Field descriptions—Define Covariates page

Field	Description
Select variables	
<Demographic variables>	(Does not apply to a logistic run that uses the HDPS option.) Analysis results are broken down by the selected variables. Demographic variables as determined by the configuration appear.

	For example, gender is typically available as a covariate. Variables appear in the analysis results as follows: • A demographic variable appears in the Cohort Summary and Data Groups tables as **<*column name*> = <*variable value*>**. • A demographic variable appears in the Coefficients and Interactions table as **<*variable value*> [<*column name*>]**.
Custom Age Group (of computed age)	(Appears if the configuration is set up to support custom age groups.) The computed age group is based on a computed age and on age groups that you define. In analysis results for all runs except logistic runs that use the HDPS option, one subset row appears for each computed age group that you define. Patients are in the age group if their computed age is within that range of the age group. A custom age group appears in the analysis results as follows: • In the Cohort Summary and Data Groups tables, **[AGEGRP(Custom)]= <*computed age group*>** appears. • In the Coefficients and Interactions table, **<*computed age group*> AGEGRP(Custom)** appears. **Notes:** • You cannot select both the **Age Group** demographic variable and the **Custom Age Group (of computed age)** option in a single run. • If the computed age is a negative number, the breakdown appears as **AGEGRP (Custom): Less than 0**. For more information, see Working with age groups.
Time category: Year	(Available for Poisson regression and logistic regression analysis runs.) Define a covariate for each calendar year, starting in January, within the analysis period defined for the analysis. In the analysis results for all runs except logistic runs that use the HDPS option, custom age groups appear as follows: • In the Cohort Summary and Data Groups tables, **YEAR = <*year*>** appears. • In the Coefficients and Interactions table, **<year> [YEAR]** appears.
One or more covariates defined as history windows	When selected, the Specify Parameters for History Window page appears later in the wizard, allowing you to create history windows. For more information, see Step 7—Specifying parameters for the history window.

Step 7—Specifying parameters for the history window

Previous step	Step 6—Defining covariates.
Description	In this step, you specify that a drug, diagnosis, or procedure occurred within a specified number of days before the index date. This range, which can be up to or including the index date, is the **history window**. For example, you might define a history window for drug history or medical history. You can define multiple history windows.

To specify parameters for a history window:

1. On the Specify Parameters for History Window page, fill in the fields according to the following table, and click **Next**.
2. To delete the history window, click **Delete**, located at the bottom of the page, near the Back and Next buttons.

   A confirmation message appears. If you confirm the deletion, the current window is deleted. If you delete all history windows, the **One or more covariates defined as history windows** checkbox on the Define Covariates page is deselected.
3. To add another history window, click **Add**, located at the bottom of the page, near the Back and Next buttons.

   The window is added immediately after the current window. The current window is saved and a new, blank Specify Parameters for History Window page appears. You can define an unlimited number of history windows.
4. Click **Next**.

   One of the following pages appears:

   - If you added another history window, a blank Specify Parameters for History Window page appears. Fill in the fields according to the following table.
   - If you did not add a history window:
     - If the analysis run is based on a logistic regression and you selected **Use high-dimensional propensity scores** on the Create Evaluative Analysis page, the Specify HDPS Parameters page appears. Continue to Step 8—Specifying HDPS parameters.
     - If the analysis run is based on the Cox proportional hazards model, the Evaluative Analysis: Cox Options page appears. Continue to Step 9—Specifying Cox options.
     - If the analysis run is based on a Poisson regression or a non-HDPS logistic regression and you have the Access Advanced Run Options permission, the Evaluative Analysis: Advanced Options page appears. Continue to Step 10—Specifying advanced options.

- If the analysis run is based on a Poisson regression or a non-HDPS logistic regression and you do not have the Access Advanced Run Options permission, the Run Options page appears. Continue to Step 11—Specifying run options.

Field descriptions—Specify Parameters for History Window page

Field	Description
Window name (prefix)	(Not applicable for a logistic analysis run that uses the HDPS option.) Name of the history window. The values that you select later on the page determine the text that appears in the SUBSET column in the analysis results. If you select values to include or exclude, the column displays the window name.
Window description	Description of the history window.
Select window variable	History variable. Available variables are determined by the configuration you are using.
Define history window as days relative to index date	Define a history window based on a start and stop day relative to the index date. • **Start**—Beginning of the history window, relative to the index date. The value must be 0 or a negative number. • **End**—End of the history window, relative to the index date. The value must be 0 or a negative number. If no index date is determined, a history window is not created.
Specify one or more inclusion values Specify zero or more exclusion values	When selected, you can create a history window to find the following information: • Patients who have any of up to 25 inclusion values during history window. • Patients who do not have any of up 25 exclusion values during the history window. • Patients who have any of up to 25 inclusion values during the window and do not have any of up to 25 exclusion values during the window. **Inclusion values** A patient is considered to have a history breakdown if any of the inclusion values are in the history window. For example, if you select DrugA, DrugB, and DrugC as co-occurring values and name the history window ABC, the analysis results contain a single row for which SUBSET is ABC, representing patients who had any of DrugA, DrugB, or DrugC during the history window. **Note:** If you select the same value that you used to define the

outcome event or index event, the value appears as a history window.

Exclusion values

You can also select exclusion values. If you select only exclusion values, a patient is considered to have the history breakdown only if none of the exclusion values are experienced by the patient in the history window. For example, if the exclusion values are DrugD, DrugE, and DrugC, the patient cannot have any of those drugs in the window.

Both inclusion and exclusion values

If you specify both inclusion values and exclusion values, a patient has the history breakdown only if the following criteria are met:

- The patient has any of the inclusion values in the window.
- The patient does not have any of the exclusion values in the window.

For example, if you specify DrugA, DrugB, and Drug C as inclusion values and DrugD, DrugE, and DrugF as exclusion values, a patient has the history breakdown if the patient has any of the following circumstances:

- DrugA during the history window and no DrugD, DrugE, or DrugF during the history window.
- DrugB during the history window and no DrugD, DrugE, or DrugF during the history window.
- DrugC during the history window and no DrugD, DrugE, or DrugF during the history window.

You cannot specify the same value as both an included value and an excluded value. However, you can include or exclude the same value that is used as the outcome event or the index event.

Selecting values

You can specify no more than 25 values. For information about selecting values, see Selecting values.

Remove terms if they are not in the source data	When selected, terms that are in the list of specified values and not in the source data are removed from the list when you continue to the next page of the wizard.
No occurrence of history term for ____ days before window start	(Always appears but is available to be selected only if either the primary index event or the comparator index event is based on the start or end of the last occurrence.) When selected, you can specify a history lead-in. The history lead-in indicates that the history window must be immediately preceded by at least the specified number of days clear of the history value and within the patient's enrollment period. If the

history lead-in is not met, a history window is not created.

The clear days occur within the patient's enrollment period, but they do not need to be within the analysis period specified for the run. For example, if the analysis period starts 100 days after the patient's enrollment start, the days can be within those 100 days before the start of the analysis period.

If history values are based on a drug variable, this option uses the specified exposure extension option when looking for days clear of the drug.

For explicitly included values, none of the selected values can have occurred for the specified number of days before the history window. For example, suppose that you specify DrugA, DrugB, and DrugC for inclusion and specify a history lead-in of 10. The history window must be preceded by at least 10 days in which there are no drug era days for DrugA, DrugB, or DrugC.

For explicitly excluded values, none of the selected values can have occurred for the specified number of days before the history window. For example, suppose that you specify DrugD, DrugE, and DrugF for exclusion and specify a history lead-in of 10. The history window must be preceded by at least 10 days in which there are no drug era days for DrugD, DrugE, or DrugF.

If you specify DrugA, DrugB, and DrugC for inclusion and DrugD, DrugE, and DrugF for exclusion, and specify a history lead-in of 10, the history window must be preceded by at least 10 days in which there are no drug era days for DrugA, DrugB, DrugC, DrugD, DrugE, or DrugF. Then, for the patient to have the history breakdown, the history window must include drug era days for any of DrugA, DrugB, or DrugC and must not include drug era days for any of DrugD, DrugE, or DrugF.

Day ranges

Some history window options are for a specified day range relative to a reference date as follows:

- A negative integer indicates days before the reference date.
- 0 indicates the reference date itself.

You can indicate a one-day history window by entering the same number as the start and end of the window.

Window start and end

When the analysis runs, the start of a history window is the later of the following options:

- Patient’s enrollment start
- Specified window start

The history window can include days that are before the analysis period start and after the patient’s enrollment start.

If the specified window start is later than either the patient enrollment end or the analysis period end, a history window is not created.

When the analysis runs, the end of a history window is the earliest of the following options:

- Enrollment end
- Analysis period end
- Specified window end

Restrictions

A restriction ensures that you cannot define a window that would start before a patient's enrollment start. The history lead-in minus the history window start cannot be greater than the lesser of the primary index lead-in and the comparator index lead-in.

Step 8—Specifying HDPS parameters

Previous step	One of the following: • Step 6—Defining covariates. • Step 7—Specifying parameters for the history window.
Relevance	Appears when you base the analysis run on a logistic regression and select **Use high-dimensional propensity scores** on the Create Evaluative Analysis page.
Description	In this step, you specify the following information: • The way that dimensions are used to create empirical variables. • The propensity score models to include. • Whether to use propensity scores to create categorical variables or matched cohorts.

To specify HDPS parameters:

1. On the Specify HDPS Parameters page, fill in the fields according to the following table, and click **Next**.
2. One of the following pages appears:
   - If you have the Access Advanced Run Options permission, the Evaluative Analysis: Advanced Options page appears. Continue to Step 10—Specifying advanced options page.
   - If you do not have the Access Advanced Run Options permission, the Run Options page appears. Continue to Step 11—Specifying run options.

Field descriptions—Specify HDPS Parameters page

Field	Description
High-dimensional variable selection	**Note:** The options in this section map to the parameters in the Pharmacoepidemiology Toolbox. The outcome_zero_cell_corr parameter is always set to 1, and other Toolbox parameters are set to the Toolbox default.
Consider ___ most prevalent values from each dimension	The analysis determines up to this number of most prevalent values for each selected dimension. These values are used to create empirical variables. The value must be an integer between 1 and 200.
Include ___ empirical covariates for propensity scores	Across all selected dimensions, up to this number of empirical variables are used in the generation of propensity scores. The value must be an integer between 1 and 500.

Ranking method	Ranking method for selecting the top *n* empirical variables, where *n* is the value specified for **Include __ empirical covariates for propensity scores**.
Infer health care service utilization	When selected, the application uses an algorithm that infers how much the healthcare service was used. Quartiles are used only if dimensions are selected and only for propensity score models that include empirical variables.
Propensity scores Demographic variables only Demographic and predefined variables Demographic, predefined, and empirical variables Demographic and empirical variables	Propensity score models to include. Each model includes different combinations of covariates, where: • Demographic variables include any demographic variables or the time category **Year** selected as covariates. • Predefined variables are history windows. • Empirical variables are the variables determined by the Pharmacoepidemiology Toolbox for selected dimensions. For more information about empirical variables, see Empirical variables.
Outcome models	
Use propensity scores to	Propensity scores are used to do one of the following options: • Create categorical variables that are deciles of propensity scores. For more information, see Categorical variables. • Create matched cohorts on the basis of propensity scores. For more information, see Matched cohorts.
Execute an unadjusted model	When selected: • If propensity scores are used to create categorical covariates, the analysis includes unadjusted results in addition to results for selected propensity score models. • If propensity scores are used to create matched cohorts, the analysis includes a run with no matching.

Step 9—Specifying Cox options

Previous step	One of the following: • Step 6—Defining covariates. • Step 7—Specifying parameters for the history window.
Relevance	Appears when you base an analysis run on a Cox proportional hazards model.
Description	In this step, you specify information that is specific to the Cox proportional hazards model.

To specify Cox options:

- On the Evaluative Analysis: Cox Options page, fill in the fields according to the following table, and click **Next**.

  One of the following pages appears:

  - If you have the Access Advanced Run Options permission, the Evaluative Analysis: Advanced Options page appears. Continue to Step 10—Specifying advanced options.
  - If you do not have the Access Advanced Run Options permission, the Run Options page appears. Continue to Step 11—Specifying run options.

Field descriptions—Evaluative Analysis: Cox Options page

Field	Description
Stratify results by	When selected, results are stratified by the covariate selected in the drop-down list. The drop-down list is populated with all the covariates you included on the Define Covariates page. If the analysis run has only one covariate, and you select **Analyze interactions** on this page, you cannot select this option.
Analyze interactions	When selected, interactions are included in the Cox proportional hazards model. If the analysis run has only one covariate, and you select **Stratify results by** on this page, you cannot select this option.
Use ___ method to handle ties	Method for handling cases with duplicate survival times.

Step 10—Specifying advanced options

Previous step	One of the following: • Step 6—Defining covariates. • Step 7—Specifying parameters for the history window. • Step 8—Specifying HDPS parameters. • Step 9—Specifying Cox options.
Relevance	Appears only if you have the Access Advanced Run Options permission.

To specify advanced options:

1. On the Evaluative Analysis: Advanced Options page, fill in the fields according to the following table, and click **Next**.

   The Run Options page appears.

2. Continue to Step 11—Specifying run options.

Field descriptions—Evaluative Analysis: Advanced Options page

Field	Description
Maximum memory available to metrics-generation engine	Maximum amount of memory to devote to the engine that generates metrics. The initial value is the lesser of following values: • The estimate of the memory required for the run. The estimate is fixed at 24.0 GB. • The maximum allowed value as determined during setup.

Step 11—Specifying run options

Previous step	One of the following: • Step 6—Defining covariates. • Step 7—Specifying parameters for the history window. • Step 8—Specify HDPS parameters. • Step 9—Specifying Cox options. • Step 10—Specifying advanced options.

To specify run options:

1. On the Run Options page, fill in the fields according to the following table, and click **Next**.

   The Name Analysis Run page appears.

2. Continue to Step 12—Naming an analysis run.

Field descriptions—Run Options page

Field	Description
Run as soon as possible Do not run until ___ at ___	When to execute the evaluative analysis run. Select **Run as soon as possible** to create the run immediately, or specify the date and time to execute the run.
Email me when complete	When selected, an email message is sent to the email address you provide after the run either completes successfully or fails. Separate multiple email addresses with commas.
Save intermediate data files	When selected, supplemental data files that are produced when a run is created are saved for viewing and downloading.

Step 12—Naming an analysis run

Previous step: Step 11—Specifying run options.

To name the analysis run:

1. On the Name Analysis Run page, fill in the fields according to the following table, and click **Next**.

   The Confirm Analysis Run Parameters page appears.

2. Review the selections for the run, and click **Submit Run**.

   A confirmation page appears, indicating that the run has been submitted. You can check the status of the run in the Status column.

   For information about results, see Viewing the results of an evaluative analysis run.

Field descriptions—Name Analysis Run page

Field	Description
Run name Description	Name and description of the run.
Add to	**Existing project**—The run is added to the project selected in the drop-down list. **New project named**—A new project is created with the name you provide, and the run is added to it.

Additional information for creating an evaluative analysis run

Empirical variables

Empirical variables are identified as follows:

1. For each selected dimension, the *x* most frequent terms are found, where *x* is the value of the **Consider __ most prevalent values from each dimension** option.

   The terms include each dimension term that occurs within the lookback period for at least one patient who is in the primary or comparator cohort and has an index date. The lookback period starts *n* days before the patient's index date and ends one day before the patient's index date, where *n* is the minimum of the primary and comparator index lead-ins that you specified when creating the run.

   For the drug dimension, a term is found for a dimension only if the drug start date is within the lookback period.

   The application looks back from the index date, with index date being considered as Day 0. For example, suppose that the index date is January 6, 2002, and the lookback period is five days. If a drug starts on January 1, 2002, the drug is found as a term for the drug dimension for the patient. If the drug starts on January 2, 2002, the is not found as a term for the drug dimension for the patient.

   The term must be within the patient's enrollment period, but not necessarily within the specified analysis period.

   Terms that are automatically or explicitly excluded are not considered.

2. For each term identified in the previous step, the following empirical binary variables are created:

   - A variable indicating that the term occurs one time.
   - A variable indicating that the term occurs sporadically.
   - A variable indicating that the term occurs frequently.

   The variable names identify the dimension, value, and frequency. For example, D01V005ONCE means dimension 1, 5th value in list, occurring one time.

   Note: A patient who is marked as having a term frequently is also marked as having the term sporadically and one time. A patient who is marked as having a term sporadically is also marked as having the term one time.

3. Using the evaluation of the binary variable and the ranking method you selected, the *y* top empirical variables across all selected dimensions are found, where *y* is the value of the **Include __ empirical covariates for propensity score** option that you specified.

Propensity scores

For each selected propensity score model, the analysis run computes propensity scores by performing logistic regression using the ridge regression option. The user-specified option to use ridge regression applies to only the final logistic regression, not to the logistic regression that generates propensity scores.

For each model, a patient is assigned a propensity score in the range of 0 to 1.

Categorical variables

If you use propensity scores to create categorical variables, the analysis run creates deciles of patients' propensity scores for each selected propensity score model. The range of scores start with the lowest score generated and end with the highest score.

The following deciles are created.

Decile	Description
0	Scores in the range of 0 to 5 percent of all the propensity scores.
1-9	Scores in a range of the next 10 percent of all the propensity scores. For example, Decile 1 is scores in the range of 5 to 15 percent.
10	Scores in the range of 95 to 100 percent of all the propensity scores.

For each selected propensity score model, a patient's propensity score can fall into a different decile.

Matched cohorts

If you use propensity scores to match cohorts, matching is performed as follows for each propensity score model:

1. Patient IDs of patients who are in the screened primary cohort and are also in N(Index) are arranged in ascending order of patient ID.

   Patient IDs for patients who are in the screened comparator cohort and are also in N(Index) are arranged in ascending order of patient ID.

2. Each group is sorted by ascending order of propensity scores. When appropriate, the order of patient IDs is retained, so if two patients have the same propensity score, they remain in order of patient ID.

3. The smaller of the following two groups is identified:

   - Patients in the screened primary cohort who are also in N(Index).
   - Patients in the screened comparator cohort who are also in N(Index).

4. For the smaller group, the application proceeds sequentially through all patients, attempting to find a patient in the other group whose propensity score matches up to 5 digits. If multiple scores in the other group are ties, the first one is used. If a match is identified, the matched pair is not considered again by the matching algorithm.

5. The previous step is repeated for matching up to 4 digits, then for up to 3 digits, then for up to 2 digits, and then for up to 1 digit.

Patients for whom a match cannot be found are not included in the analysis for that propensity score model.

Viewing the results of an evaluative analysis run

Viewing the results of an evaluative analysis run

Viewing the results of an evaluative analysis run

Prerequisites

- You must have the View Analysis Results permission.

To view the results of an evaluative analysis run:

1. Select the **Evaluative Analysis** tab.
2. To filter the analysis results that appear, select options from the **Project** and **Configuration** drop-down lists.
3. Navigate to the row containing the run for which to view results.

   Note: The run must have completed successfully. To verify that the run has completed, check the **Status** column.

4. Select the row menu (), and select **View Results**.

   The results appear. Multiple types of results are available.

5. To view each of the results, click a link below the Evaluative Analysis tab. The run type determines the options that are available:

   - Cohort Summary
   - Coefficients
   - Interactions
   - Data Groups
   - Survival Graph
   - Survival Comparison Graph

   Several columns include patient counts with links that allow you view details about the patients. For more information, see About drilldown.

6. To view notes about the run, click **Show Notes**. Hide the notes by clicking **Hide Notes**.

Cohort Summary results

For information about how to view this page, see Viewing the results of an evaluative analysis run.

A row does not appear for a non-ALL subset with an N(Index) value of 0.

Field descriptions—Cohort Summary page

Field	Description for the ALL subset	Description for other subsets, if different from the ALL subset
Index Term	Name of the index event and an indication of whether it is the primary or comparator index event.	
N(Incidence)	Number of patients who have the outcome event and are also in the N(Index) population. **If the outcome event is a diagnosis or procedure** Number of patients who have that diagnosis or procedure within follow-up. **If the outcome event is a drug** Number of patients who have the start of a drug era for that drug within follow-up. If a drug era starts before follow-up and continues into follow-up, the patient is not counted as having the outcome event during follow-up unless the patient has a separate, subsequent drug era that starts within follow-up. This rule applies regardless of whether an exposure extension was specified for the outcome event. For example, suppose that the follow-up start is October 1, 2005 and a patient has only two prescriptions for the drug: • September 28, 2005, 10-day supply. • October 3, 2005, 10-day supply. These prescriptions make up one drug era that is from September 28,	Number of patients who are counted in N(Incidence) for the ALL subset and who meet the following criteria: **For breakdown by year** Patients whose outcome event during follow-up is within the year. Only the start of a drug era is considered. **For history window breakdown** Patients have the history value in the history window. **For demographic breakdown** Patients have the demographic characteristic.

	2005 through October 12, 2005. The patient is not counted in N(Incidence) because the patient does not have the start of a drug era during follow-up. If a separate era started during follow-up, the patient would be counted in N(Incidence).	
N(Index)	Number of patients with an index date.	Number of patients who are counted in N(Index) for the ALL subset and who meet the following criteria: **For breakdown by year** Patients have any part of follow-up within the year. This count uses follow-up days, which might start later than the index date. For example, if the index is in 2004 and follow-up start is in 2005, the patient is not in N(Index) for 2004 because follow-up does not start until 2005. Follow-up end depends on how follow-up was defined and whether the outcome event occurs during follow-up. **For history window breakdown** The patient has the history value in the history window. **For demographic breakdown** Patients have the demographic characteristic.
N(Initial)	Number of patients who are in the screened cohort and have at least one day of enrollment during the specified analysis period. **Note:** The **View Excluded Patients** option is available from the menu in the N(Initial) column for the ALL subset. When you select this option, patients who meet either of the following criteria appear: • Patients were screened out of the cohort. • Patients are in the screened cohort but have no days of enrollment within the analysis period.	Number of patients who are counted in N(Initial) for the ALL subset and meet the following criteria: **For breakdown by year** Patients have any enrollment days within the year. **For history window breakdown** This column is always empty. **For demographic breakdowns** Patients have the demographic characteristic.

Person Years	(Available only for Poisson and Cox analysis runs.) Total number of person years in follow-up. The following calculation is used: person-days / 365.24 Follow-up start and end depend on how follow-up was defined for the analysis. Person Years are from follow-up start up to and including the last day of follow-up. For patients counted in N(Incidence), the last day of follow-up is the date of the outcome event. You select follow-up options when you specify temporal parameters.	Total number of person years in follow-up that are counted for the ALL subset and that meet the following criteria: **For breakdown by year** Person-days are within the year. **For history window breakdown** Person years are for patients who have the history value in the history window. **For demographic breakdown** Person years are for patients who have the demographic characteristic.
SUBSET	The following rows appear in the table: • One row for which ALL appears in the SUBSET column, and the index term is the primary index term. • One row for which ALL appears in the SUBSET column, and the index term is the comparator index term. • One row for each combination of covariates (or each decile for an HDPS run that uses deciles) that has at least one patient counted in the N(Index) column. For an HDPS run that uses deciles, there is one row for each decile that has at least one patient counted in the N(Index) column. If the run uses matched cohort, the ALL row is for an analysis that does not use matched cohorts.	

Coefficients results

For information about how to view this page, see Viewing the results of an evaluative analysis run.

On the Coefficients table, you view statistics for individual predictors.

Different fields appear in the results, depending on the run type.

If a predictor has only two possible values, only the value with the lowest patient count for N(Index) is shown as a predictor.

If a predictor has more than two possible values, all the values appear as rows but the predictor with the highest patient count for N(Index) has a value of 0 in the Coeff field and a value of 1 in the RR field.

Field descriptions—Coefficients page for a Poisson analysis run

Field	Description
Coeff	Coefficients are generated as estimated regression model parameters from a model of the form Y = a + Xb, where: • Y is the vector of fitted predictions on the log scale. N(Incidence) ~ Poisson with mean (Person Years) * eY. • a is a constant (intercept). • b is a vector of regression parameters. • X is a design matrix. The model design matrix uses binary 0-1 predictor columns with 0 assigned to the group with the largest N(Index). Therefore, coefficients represent the natural log of the risk ratio of risk between the specified group and the group with the largest N(Index). If a covariate has more than two possible values, the predictor with the highest patient count for N(Index) is used as the comparator, so the Coeff value and SE columns are 0 for that row.
IR	Raw incidence rate. The following calculation is used: N(Incidence) / Person Years
N(Incidence)	Number of patients with the outcome event during follow-up.
N(Index)	Number of patients with an index date.
Person Years	Number of person years in follow-up.
Predictor	This value includes the following information:

	• <intercept> • Primary index term Additional predictors are the values of the covariates selected on the Define Covariates page of the run wizard.
Response	Term specified as the outcome event.
RR	Fitted relative risk. The following calculation is used: exp(Coeff)
RR_025	Lower 2.5% confidence limit for RR.
RR_975	Upper 97.5% confidence limit for RR.
SE	Standard error of the value in the Coeff field.
Z Score	The following calculation is used: Coeff / SE

Field descriptions—Coefficients page for a logistic analysis run, with or without HDPS

Field	Description
Coeff	Coefficients are generated as estimated regression model parameters from a model of the form Y = a + Xb, where: • Y is the vector of fitted predictions on the logistic scale. • N(Incidence) ~ Binomial[N_Index, p(Y)], where p(Y) = Fitted Prob = 1/(1 + e-Y). • a is a constant (intercept). • b is a vector of regression parameters. • X is a design matrix. The model design matrix uses binary 0-1 predictor columns with 0 assigned to the group with the largest N(Index). Therefore, coefficients represent the natural log of the odds ratio of risk between the specified group and the group with the largest N(Index). If a covariate has more than two possible values, the predictor with the highest patient count for N(Index) is used as the comparator, so the Coeff value and SE columns are 0 for that row.
Model	(Available only for an HDPS run.) Propensity score model.
N(Incidence)	Number of patients with the outcome event during follow-up.

N(Index)	Number of patients with the index event. If the analysis run has a year subset, the number of patients with any part of follow-up during the year.
OR	Odds ratio for the outcome occurring during follow-up. The following calculation is used: exp(Coeff)
OR_025	Lower 2.5% confidence limit for OR.
OR_975	Upper 97.5% confidence limit for OR.
Predictor	For non-HDPS logistic analysis runs, the predictors are <intercept>, the primary index term, and each covariate. For HDPS logistic analysis runs, if propensity scores were used to create categorical variables, the predictors are deciles of propensity scores, in addition to <intercept> and the primary index term. If propensity scores were used to match cohorts, the predictors are <intercept> and the primary index term.
Proportion	The following calculation is used: N(Incidence) / N(Index)
Response	Term specified as the outcome event.
SE	Standard error of the value in the Coeff field.
Z Score	The following calculation is used: Coeff / SE

Field descriptions—Coefficients page for a Cox analysis run

- If **Analysis interactions** was not selected for the analysis run, a single Coefficients table appears.
- If **Analysis interactions** was selected, multiple Coefficients tables appear.

If a stratification variable was used for the analysis run, the variable does not appear as a predictor.

Field	Description
Coeff	Regression coefficient.
P Value	Two-sided P-value based on the normal distribution for the Z Score.
Predictor	Primary index term or covariate value.

Response	Outcome event.
RR	Fitted relative risk. The following calculation is used: exp(Coeff)
RR_025	Lower 2.5% confidence limit for RR.
RR_975	Upper 97.5% confidence limit for RR.
SE	Standard error of the value in the Coeff field.
Z Score	The following calculation is used: Coeff / SE

Interactions results

For information about how to view this page, see Viewing the results of an evaluative analysis run.

Field descriptions—Interactions page

Field	Description
Fitted N(Incidence)	Number of patients expected to have the outcome event during follow-up. • The following calculation is used for a Poisson run: (Person Years * Fitted IR) summed over all patients having the two predictor values • The following calculation is used for a logistic run, with or without the HDPS option: Fitted Prob summed over all N(Index) patients having the two predictor values
N(Incidence)	Number of patients with the outcome event during follow-up who have both predictor values.
N(Index)	Number of patients with the index date who have both predictor values.
Person Years	Number of person years in follow-up.
Predictor 1 Predictor 2	Covariate value or primary index event.
Response	Outcome event.
Z Score	Lack of fit statistic. The following calculation is used: [(N(Incidence) – Fitted N(Incidence)] / Standard Error of [(N(Incidence) – Fitted N(Incidence)]

Data Groups results

For information about how to view this page, see Viewing the results of an evaluative analysis run.

The Data Groups table has one row for each combination of index term and subset (and, for HDPS runs, outcome model type) that appears in the Cohort Summary. ALL subset rows do not appear.

A row does not appear for a non-ALL subset with an N(Index) value of 0.

Field descriptions—Data Groups page for a Poisson analysis run

Note: Additional columns are available for the individual covariate variables and values that are in the SUBSET column.

Field	Description
Fitted IR	Fitted incidence rate. The following calculation is used: Fitted N(Incidence) / Person Years
Fitted N(Incidence)	Number of patients expected to have the outcome event during follow-up, according to the fitted model.
Index Term	Name of the index event and whether it is the primary or comparator index event.
N(Incidence)	Number of patients with the outcome event during follow-up.
N(Index)	Number of patients with the index date.
Person Years	Number of person years in follow-up.
Response	Outcome event.
SUBSET	Combination of covariates, such as demographics, year, or history window.
Z Score	Lack of fit statistic. The following calculation is used: [(N(Incidence) – Fitted N(Incidence)] / Standard Error of [(N(Incidence) – Fitted N(Incidence)]

Field descriptions—Data Groups page for a non-HDPS logistic analysis run

Note: Additional columns are available for the individual covariate variables and values that are in the SUBSET column.

Field	Description
Fitted N(Incidence)	Patients expected to have the outcome event within follow-up. The following calculation is used: [N(Incidence) + N(NoIncidence)] * Fitted Prob
Fitted Prob	Fitted probability that the patient will have outcome event during follow-up.
Index Term	Name of the index event and whether it is the primary or comparator index event.
N(Incidence)	Number of patients with the outcome event during follow-up.
N(NoIncidence)	Number of patients who do not have outcome event during follow-up.
Prob Pctile	Percentile rank of the value in the Fitted Prob column.
Response	Outcome event.
SUBSET	Combination of covariates, such as demographics, year, or history window.
Z Score	Lack of fit statistic. The following calculation is used: [(N(Incidence) – Fitted N(Incidence)] / Standard Error of [(N(Incidence) – Fitted N(Incidence)]

Field descriptions—Data Groups page for an HDPS logistic analysis run

Field	Description
Fitted N(Incidence)	Patients expected to have the outcome event within follow-up. The following calculation is used: [N(Incidence) + N(NoIncidence)] * Fitted Prob
Fitted Prob	Fitted probability that the patient will have outcome event during follow-up.
Index Term	Name of the index event and whether it is the primary or comparator index event.
Model	Propensity score model.

N(Incidence)	Number of patients with the outcome event during follow-up.
N(NoIncidence)	Number of patients who do not have outcome event during follow-up.
Prob Pctile	Percentile rank of the value in the Fitted Prob column.
Response	Outcome event.
Subset	Combination of covariates, such as demographics, year, or history window.
Z Score	Lack of fit statistic. The following calculation is used: [(N(Incidence) – Fitted N(Incidence)] / Standard Error of [(N(Incidence) – Fitted N(Incidence)]

Survival Graph results

For information about how to view this page, see Viewing the results of an evaluative analysis run.

A Survival Graph, which appears for a Cox analysis run, uses Kaplan-Meier survival estimates to show the probability that a patient will not have the outcome event (that is, will survive) over time, where time is represented as person-years of follow-up.

A note in the graph indicates the predictors that the curves are based on. The predictors are those that have the highest counts for N(Index).

Covariates appear as Covariate_1, Covariate_2, and so on. A key below the graph indicates the actual names of the covariates.

When a stratification variable is selected for the run

- A line appears in the graph for each value of the stratification variable. Each line is for the combination of the screened primary cohort and the screened comparator cohort.
- No lines appear for confidence limits.

When a stratification variable is not selected for the run

- One line appears in the graph for the combined primary and comparator cohorts.
- Lines appear for the upper and lower confidence limits.

Survival Comparison Graph results

For information about how to view this page, see Viewing the results of an evaluative analysis run.

A Survival Comparison Graph, which appears for a Cox analysis run, plots the screened primary cohort and screened comparator cohort separately. The graph uses Kaplan-Meier survival estimates to show the probability that a patient will not have the outcome event (that is, will survive) over time, where time is represented as person-years of follow-up.

Two lines appear in the graph, one each for the primary index event and comparator index event.

A note on the graph identifies the following information:

- For covariates not used for stratification, the value with the highest count for N(Index).
- For the covariates used for stratification, if any exist, the value with most distinct values of person-years in follow-up.

Covariates appear as Covariate_1, Covariate_2, and so on. A key below the graph indicates the actual names of the covariates.

Creating and running a definition file

For an existing analysis run, you can create a definition file that allows you to submit the analysis run automatically instead of through the Empirica Healthcare Analysis user interface. You can run the definition file as often as needed.

Prerequisites

- You must be a superuser.
- The analysis run must have completed.

To create and run a definition file:

1. Select one of the following tabs:
   - Descriptive Analysis
   - Evaluative Analysis
   - Analysis Setup
2. Select the row menu () for an analysis run, and select **Create Definition File**.

   The Run Definition File dialog box appears and contains the contents of the file.
3. Select the full text in the dialog box.
4. Paste the text into a text application, such as Notepad.
5. To modify the name of the analysis run, edit the **name** parameter.

   Note: Do not modify the values of any other parameters.
6. Save the document with an extension of **IN**, and close the document.
7. Copy the file to the correct directory:
   A. Open the webvdme.properties file, located in the following directory on the application server:

      $INSTALL_DIR/Healthcare/WEB-INF/classes

   B. Find the temp_dir parameter, and note the directory it references, such as:

      temp_dir=/u01/app/oracle/product/EHC/temp

   C. Navigate to the specified directory, and locate the **input** folder.

      For example, for the previous example, the directory is

      /u01/app/oracle/product/EHC/temp/input

   D. Place the IN file in the input directory.

The listener process checks the input directory every 2 to 30 seconds and executes any run definition files with the IN extension. While being processed, the file has an extension of PROC. The file extension changes to DONE when the execution of the run definition file is complete.

Publishing an analysis run

You can publish any run that you create. When you publish a run, the results are available to other users, including those who might not otherwise have permissions to view the results and those who might not have privileges to create runs and use the configuration. Publishing allows multiple users to review the results of data that is analyzed only a single time, reducing machine time overhead.

You can publish a run as soon as it is submitted.

1. Select the **Descriptive Analysis** or **Evaluative Analysis** tab.

2. Select the row menu () for a run, and select **Publish**.

Patient Queries tab

Patient Queries tab

On the Patient Queries tab, you can define and save patient groups, and run queries based on the specified queries. The Patient Queries tab has the following views:

- Patient Groups (default view)
- Query Library

About queries

A query is a set of criteria that you specify based on variables and values in the source data to search for a set of patients in the application. In the Query Library view, use the Query Wizard to create or edit a query. The Query Wizard guides you through the steps of selecting variables, selecting values, and specifying query logic.

A query is applied to the data at the time you run a query. If the data changes after you run a query, the query results might differ when you re-run the query.

When creating a patient group from the results of a query , you can create new queries or use existing queries. The list of existing queries is referred to as the Query Library.

Note: Queries do not include custom terms defined during the creation of analysis runs, because they are run against source data instead of analysis results.

Patient Groups view

Patient Groups view

In the Patient Groups view, you can view a list of all patient groups. A patient group is a collection of patients who meet the specified set of criteria. You can then run an analysis or a report against the patient group. For example, you might want to save a list of all patients who fit a particular demographic profile and who have taken a particular class of drugs.

You can drill down on the list of patients to view their details.

After you create a patient group, you can manually add or transfer additional patients to the patient group.

You can access the details of individual patients in a patient group. You can review a patient group and mark patients for exclusion. Patients marked as excluded are not included in reports.

General activities

The following links and filters appear at the top of the page and affect the entire page.

- Create Empty Patient Group
- Columns
- Print
- Download
- Select Rows
- Filter by Project or Configuration

Row-specific activities

The following menu options are available from the row menu, located in the left-most column of the table, and affect an individual row in the table:

- View Patients
- View Query
- Rename
- Report—Generate a report for the patient group.
- Publish—Make the patient groups that you have created available for other users.
- Add Query to Library—Available for patient groups created from the Query Library view.
- Copy
- Delete

Field descriptions—Patient Groups view

Field	Description
# of Patients	Number of patients in the patient group.
Associated Run	When a patient group is created from a run result, then it contains one of the following values: • Name of the analysis run that generated the results. If the run associated with a patient group is deleted, the patient group is retained but is no longer associated with the run, and the run name is removed from the column. • Empty—If the patient group was not created from the results of an analysis run.
Configuration	Name of the configuration associated with the patient group.
Created	Date and time when the patient group was created.
Created By	User name of the person who created the patient group.
Description	Description of the patient group.
ID	Identifier that was assigned to the patient group when it was created. The patient group IDs are unique and are not reused if they are deleted.
Modified	Date and time when the patient group was last modified.
Modified By	User name of the person who modified the patient group.
Name	Name of the patient group.
Patients Added	• **NA (Not Applicable)**—The patient group was not created using the Create Using Query Wizard option from the Query Library view. • **Yes**—The patient group was created using the Create Using Query Wizard option, and the patients were manually added to or transferred to the patient group. • **No**— The patient group was created using the Create Using Query Wizard option, and no patients were added or transferred to the patient group.

Status	Status of patient group processing.
	• If the patient group is created as a background job, the column is empty until the job is complete. Then, one of the following values appears:
	  • **Completed—**The job completed successfully.
	  • **Error Occurred—**The job failed, resulting in a failed patient group.
	  • **Cancelled—**Creation of the patient group was cancelled.
	  While a background job for a patient group is running, you can view the job status by clicking the progress indicator () at the top of the page.
	• If the patient group is not created as a background job, then the following value appears:
	  • **Completed—**The job completed successfully.

Viewing a query of a query-based patient group

You can view the query, if a patient group is created using the query wizard from the Query Library view.

1. Select the **Patient Queries** tab.

   The Patient Groups page appears.

2. Select the row menu (), and click **View Query.**

   A dialog box appears, listing the variables and values that make up the query as well as the query logic.

Adding a query to the query library

You can create a query from the patient group, if a patient group is created using the query wizard from the Query Library view. With this option you can re-create a deleted query.

1. Select the **Patient Queries** tab.

   The Patient Groups page appears.

2. Select the row menu (), and click **Add Query to Library.**

   The query is added to the query library and appears on the Query Library page.

Combining patient groups

Prerequisites—You must have the Create Queries/Patient Queries permission.

When you combine two or more patient groups into a new patient group, the Reviewed and Excluded statuses for all patients in the new patient group are null.

Query Library View

Query Library view

In the Query Library view of the Patient Queries tab, you can view a list of all the queries. A query is the saved specification of criteria for finding patients. You can create a query for any configuration to retrieve a list of patients who meet the criteria. For example, you can create a query to find all female patients who are between the ages of 17 and 45 and who have taken DrugA.

You can save and reuse queries to retrieve patients of interest as needed.

General activities

The following links and filters appear at the top of the page and affect the entire page.

- Create Using Query Wizard
- Columns
- Print
- Download
- Select Rows
- Filter by Project, Configuration, or Origin

Row-specific activities

The following menu options are available from the row menu, located in the left-most column of the table, and affect an individual row in the table:

- Create Patient Group
- Run
- View
- Edit
- Rename
- Report
- Publish
- Copy
- Delete

Field descriptions—Query Library view

Field	Description
Configuration	Name of the configuration associated with the query.
Created	Date and time when the patient group was created.
Created By	User name of the person who created the patient group.

Description	Description of the query.
ID	Identifier that was assigned to the query when it was created. The query IDs are unique and are not reused if they are deleted.
Modified	Date and time when the patient group was last modified.
Modified By	User name of the person who modified the patient group.
Name	Name of the query.
Origin	Context in which the query was originally created: • **Query**—The query was created from the Query Library page. • **Patient Group—**The query was created by using the **Add Query to Library** option for a patient group on the Patient Groups view.
Project	Name of the project associated with the query.

Creating a query

Creating a query using the Query Wizard

A query consists of one or more SQL clauses linked by logic into a SQL statement. When creating a query, you must specify a configuration for it. However, you can rerun a query on a different compatible configuration at any time. Creating a query consists of the following steps:

1. Specify a configuration.
2. Select variables and values, and define the logic for the SQL query.
3. Review the SQL statement and its logic.
4. Name or save the query

Step 1—Specifying a configuration

1. Select the **Patient Queries** tab.
2. Click **Query Library**.

   The Query Library page appears.
3. Click **Create Using Query Wizard**.

   The Select Configuration page appears.
4. Select a configuration from the drop-down list, or click **Browse** to navigate to a configuration, and click **Next.**

   The Define Query page appears.
5. Continue to Step 2a—Specifying variables and values.

Step 2a—Specifying variables and values

For variables in a query, values do not need to be predefined. You can provide values when the query is run. For example, you can create a query to find patients who are using two drugs, where you specify both the drugs (DrugA and DrugB) when you run the query.

Previous step	Step 1—Specifying a configuration.
Description	In this step, you select variables and their values for the configuration selected in the previous step. Later in the wizard, you define the logic for the query.

To specify variables and values:

1. On the Define Query page, select variables from the source data:

   A. Click **Select Variables**.

      The Select Variables dialog box appears.

   B. Use the arrow buttons to move variables from the All Variables in All Tables list to the Selected Variables list.

      Note: You can include a variable in a query multiple times. For example, suppose you want to include patients who are between 17 and 25 or who are above 65. Include the Age variable two times, with the first instance set to 17-25 and the second instance set to above 65. Connect the variables with the OR operator.

   C. When you have finished, click **OK**.

      One panel appears for each variable.

2. Select values for each variable. The way that you specify values depends on the way the variable was set up in the configuration. Methods for specifying values include the following:

   - Select values from a hierarchy.
   - Select values from a list of variables.
   - Specify a range of values.
   - Specify free text.

3. If the operators in the query are all AND, all OR, all INTERSECT, or all UNION, you do not need to specify values for every variable. If you do not specify a value for a variable and you do not check **Include Null values** for the variable, the variable is ignored. If all the operators in the query are MINUS or have a combination of different operators, you must specify values for every variable.

4. To include null (empty) values along with the other values that you specify for a variable, select **Include Null values**. For more information, including how the NOT operator works if you include null values, see Defining the query logic.

5. To remove a condition, click **[X]** in the upper-right corner of the panel containing the variable. Every condition that you specify must be referenced by the query logic.

6. Continue to Step 2b—Defining query logic.

Step 2b—Defining query logic

In the conditions of a query, the variables are connected by the AND operator by default. You can modify the operators as needed. When you specify query logic, refer to the conditions using the numbers that were assigned automatically to the variables. For example, the query logic **1 AND 2** means **Condition 1 AND Condition 2**.

Previous step	Step 2a—Specifying variables and values.
Description	In this step, you define the query logic that connects the variables selected in the previous step. Later in the wizard, you review the SQL statement and its logic.

To change the query logic string:

1. At the bottom of the Define Query page, click **Edit**.

   The Edit Logic dialog box appears.

2. Modify the logic using the logical operators, set operators, and parentheses as needed.

   You can use the following types of operators in a logic string:

   - Logical operators (AND, OR, and NOT).
   - Set operators (INTERSECT, UNION, and MINUS).

   For more information on SQL logics refer to SQL operators.

3. When you are finished editing the logic, click **OK**.

4. Click **Next.**

   The Preview Query page appears.

5. Continue to Step 3—Reviewing the SQL statement and its logic.

Step 3—Reviewing the SQL statement and its logic

Previous step	Step 2b—Defining query logic.
Description	In this step, you review the defined SQL statement and its logic. Later in the wizard, you name the query and save it.

1. On the Preview Query page, review the SQL statement and its logic. Click **Back** to make changes, or click **Next** to advance to the next page.

   The Save Query page appears.

   Note: If query values were not specified for a variable, **?** shows as the value for the variable.

2. Continue to Step 4—Saving the query.

Step 4—Saving the query

Previous step	Step 3—Reviewing the SQL statement and its logic.
Description	In this step, you name the query and save it.

1. On the Save Query page, type a **Name** and **Description.**
2. Select a project to be associated with the new query:
   - **Add to existing project—**Select an existing project from the drop-down list
   - **Add to new project named**—Type the **Name** of a new project.
3. Click **OK**.

   The query is saved and appears in the Query Library.

Creating a patient group

Creating a patient group

You can create a patient group from the following locations:

- Patient Groups tab.
- Queries tab.
- Pages that list patient IDs or counts of patient IDs (for example, analysis run results or counts in the Browse Data dialog box).

When you select Create Patient Group option for a query, the patient group is created as a background job. But when you click Run, the patient group is created while you wait and as the results are displayed, you can create a patient group from them.

Note: Oracle recommends that you run queries as background jobs.

You have the following options for creating patient groups:

- Create an empty patient group—Create an empty patient group and provide an explicit list of patient identifiers by entering them or by transferring patients to the patient group.
- Create a patient group from query results—Create a query that retrieves a list of patient identifiers according to a set of specified criteria using SQL.

Creating an empty patient group

After you create an empty patient group, you can add patients to the group.

1. Select the **Patient Queries** tab.

   The Patient Groups page appears.

2. Click **Create Empty Patient Group**.

   The Create Empty Patient Group page appears.

3. Select a configuration from the drop-down list, or click **Browse** to navigate to a configuration, and click **Next**.

4. Fill in the fields, and click **Save**.

   An empty patient group is created.

5. View the empty patient group and add patients to it using one of the following options:

   - Manually add patients.
   - Transfer patients.

Manually adding patients to a patient group

You can add patients manually to any patient group (empty or non-empty) that you have created or that has been published to your login group.

1. Select the **Patient Queries** tab.

   The Patient Groups page appears.

2. Select the row menu (), and click **View Patients**.

   The Patients page appears.

3. Click **Manually Enter IDs**.

   The Append Manually Entered Patient IDs dialog box appears.

4. Type or paste a list of patient IDs separated by commas.

5. Click **Append**.

   The application searches the source data for patients with the specified patient IDs. When search is complete, the following information appears on the page:

   - The number of patient IDs that were entered manually.
   - The number of patient IDs that were actually added to the patient group.
   - The total number of patients now included in the patient group.

   If a patient ID does not exist in the source data, that patient ID is not added to the patient group.

6. Click **Close**.

Note: For a query-based patient group, Oracle recommends that you do not add patients that do not meet query criteria, even though the application allows it. If the query is executed again, the added patients are not be included because they do not meet the query criteria.

Transferring patients to a patient group

You can transfer patients to an existing patient group from various pages of the application or from application reports that display a list of patients identified by the patient ID.

1. Do one of the following:

   - When viewing a list of patients, click the patient, and from the drop-down menu, select **Transfer to Patient Group**.
   - When viewing a table, click a patient count, and from the drop-down menu, select **Transfer to Patient Group**.
   - When viewing a graph, click an element of the graph, and from the drop-down menu, select **Transfer to Patient Group**.

   The Transfer to Patient Group dialog box appears.

2. Select a patient group, and click **OK**.

   You can select a patient group from the list of patient groups available on the basis of the data configurations that you are viewing.

   A message with number of patients transferred to the patient group appears.

Note: For a query-based patient group, Oracle recommends that you do not add patients that do not meet query criteria, even though the application allows it. If the query is executed again, the transferred patients would not be included because they would not meet the query criteria.

Creating a patient group from query results

This option runs as a background job and is recommended if you are working with large sets of patients.

To run and create a patient group from the query results:

1. Select the **Patient Queries** tab.

   The Patient Groups page appears.

2. Select the **Query Library** view, located just below the Patient Queries tab.

   The Query Library page appears.

3. Select the row menu (), and click **Create Patient Group**.

   The Select Configuration page appears.

4. Select a configuration from the drop-down list, or click **Browse** to navigate to a configuration, and click **Next.**

   The Run Query page appears.

5. Review variables and values. If values are not defined for the variables in the query, define them and click **Next**.

   The Create Patient Group page appears.

   Note: If the operators in the query are all AND, all OR, all INTERSECT, or all UNION, you do not need to specify values for every variable. If you do not specify a value for a variable and you do not check **Include Null values** for the variable, the variable is ignored. If all the operators in the query are MINUS or have a combination of different operators, you must specify values for every variable.

6. Fill in the fields, and click **OK**.

SQL operators

SQL operators

In the Empirica Healthcare Analysis application, you can use the following types of SQL operators to modify query logic for patient queries.

- Logical operators
- Set operators

Logical operators

SQL logical operators act on individual rows of the source data tables. You can use the following logical operators between conditions:

- **AND**—Find patients for which both conditions occurred simultaneously.
- **OR**—Find patients for which either of the conditions occurred.
- **NOT**—Find patients that do not meet the condition. The logical operator NOT is used to negate a condition. You cannot use NOT alone to connect conditions.

Logical operators cannot act on the results of set operators.

For information on operator priorities, see Operator priority.

Set operators

SQL set operators act on sets of patients that are retrieved by the conditions on each side of the set operator. You can use the following set operators between conditions:

- **INTERSECT**—Find patients that are in both sets.
- **UNION**—Find patients that are in either of the sets.
- **MINUS**—Find patients that are in the set to the left of the operator, and subtract from that list the patients that are in the set to the right of the operator.

For some queries, you can specify the same query using either logical operators or set operators. In those situations, Oracle recommends using logical operators because they are generally more efficient than set operators.

For information on operator priorities, see Operator priority.

Operator priority

Operators are applied in the following order when the query expression is interpreted:

- NOT
- AND
- OR
- INTERSECT, UNION, MINUS (same priority)

The logical operators cannot perform any action on the result of the set operators. For example, the following query is not valid:

1 AND (2 INTERSECT 3)

You can use parentheses to change the order in which a query expression is interpreted. If you do not use parentheses explicitly, the query is interpreted as if there are parentheses, based on the default order of operators.

For example:

1 AND 2 OR 3 INTERSECT 4

With no supplied parentheses, the Empirica Healthcare Analysis application reads the query as follows:

((1 AND 2) OR 3) INTERSECT 4

If you use parentheses as follows, the result of the query will be different from the result of query with no parentheses:

1 AND (2 OR 3) INTERSECT 4

Examples of query logic

Suppose that you specify the following conditions:

1. Condition 1—Drug: Thiamine
2. Condition 2—Drug: Niacin
3. Condition 3—Dose: From 20—Dose is greater than or equal to 20.

For illustrative purposes in the following examples, the query logic does not always refer to all the three conditions. When you are creating a query in the Empirica Healthcare Analysis application, the query logic must refer to all the conditions in the query.

Besides, a condition can also refer to multiple values. For example, you could include the Drug variable one time and select Thiamine and Niacin as one condition. The Drug variable is included two times in this example so that complex logic can be illustrated.

Suppose that the source data contains patients A through G with the following data:

PATIENT_ID	DRUG	DOSE
A	Thiamine	25
A	Niacin	15
B	Thiamine	15
B	Niacin	25
C	Calcium	15
D	Thiamine	15
E	Thiamine	25
F	Niacin	25
G	Calcium	25

Query logic for logical operators:

- **1 AND 3**

  The query retrieves patients for whom Drug Thiamine was administered, and Dose is greater than or equal to 20.

  A

  E

- **1 OR 3**

  The query retrieves patients for whom either Drug Thiamine was administered, or Dose is greater than or equal to 20.

  A

  B

  D

  E

  F

G

If the OR operator is used between variables from different tables, the query displays the patients who are in each of those tables. For example, suppose a query specifies that:

"Death Date from the Demog table is within a specified range" **OR** "Event from the Event table is Death" **OR** "Disposition from the Proc table is Died".

The query result contains only those patients who meet the criteria and are in all three tables.

You may also use set operators instead of logical operators. For this example, replacing OR with UNION in the query above gives the same result.

- **1 AND 2**

  No patients are retrieved because a single row in the Drug table cannot have both Thiamine and Niacin for a Drug. There must be two separate rows for each drug.

  You can use the set operator INTERSECT to find patients for whom both Thiamine and Niacin were administered. See the following 1 INTERSECT 2 example.

- **1 AND 2 OR 3**

  The query retrieves patients for whom Drugs Thiamine and Niacin were administered (not possible, because a row can have only one Drug), or Dose is greater than or equal to 20.

  A

  B

  E

  F

  G

  You can re-write this query logic as **(1 AND 2) OR 3**.

- **1 AND (2 OR 3)**

  The query retrieves patients for whom Drug Thiamine was administered, and for the same row, either Drug Niacin was administered (not possible, because a row can have only one Drug) or Dose is greater than or equal to 20.

  A

  E

- **1 AND NOT 3**

  The query retrieves patients for whom Drug Thiamine was administered and Dose is not greater than or equal to 20.

  B

  D

- **NOT 1 AND 3**

  The query retrieves patients for whom Drug Thiamine was not administered and Dose is greater than or equal to 20.

  B

  F

G

You can re-write this query logic as **(NOT 1) AND 3**.

- **NOT (1 AND 3)**

  The query retrieves patients for whom Drug Thiamine with Dose greater than or equal to 20 was not administered.

  A

  B

  C

  D

  F

  G

- **NOT 1 AND NOT 3**

  The query retrieves patients for whom Drug Thiamine was not administered, and Dose is not greater than or equal to 20 for that.

  A

  C

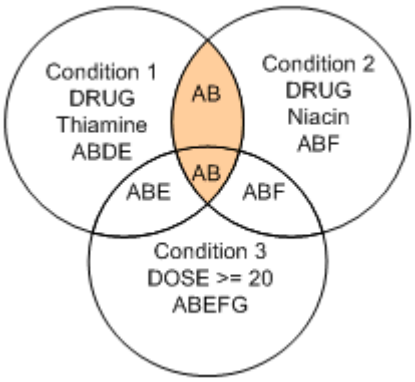
Query logic for set operators:

- **1 INTERSECT 2**

  The query results in the intersection of the two sets where Set 1 has all the rows where Drug Thiamine was administered, Set 2 has all the rows where Drug Niacin as administered, and the intersection is all the same patients appearing in both Set 1 and Set 2.

  A

  B



- **1 INTERSECT 3**

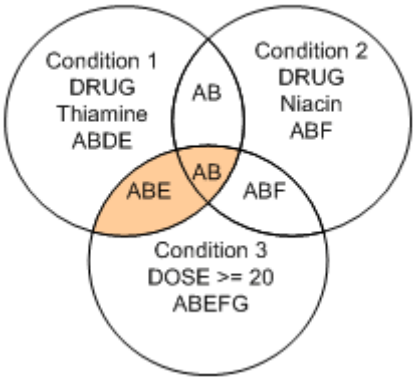
  The query results in the intersection of the two sets.

  A

  B

  E

Here, for patient B, Drug Thiamine was not administered where Dose is greater than equal to 20, but patient B is in the set of Condition 3, as well as in the set of Condition 1, therefore, the query retrieves patient B.



- **1 UNION 2**

  The query results in the union of the two sets.

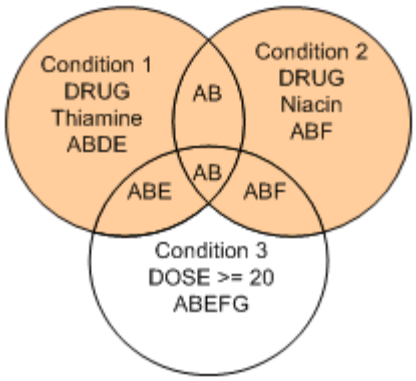
  A

  B

  D

  E

  F



- **3 MINUS 1**

  The query results in the patients remaining after patients in Set 1 are removed from Set 3. Set 3 has all the patients with a Dose of greater than 20, and Set 1 has all the patients for whom Thiamin was administered. So 3 MINUS 1 finds all the patients in Set 3 except for whom Drug Thiamine was administered.

  F

  G

- **1 INTERSECT 2 AND 3**

  The query results in the following patient:

  B

  You can re-write this query logic as **1 INTERSECT (2 AND 3)**.

- **1 UNION 2 AND 3**

  The query results in the following patients:

  A

  B

  D

  E

  F

  You can re-write this query logic as **1 UNION (2 AND 3)**.

Null values:

If you select **Include Null values** for a condition, the condition finds patients for which the condition is true or the variable referenced by the condition is null. If the condition is preceded with the NOT operator, the condition finds patients with rows for which the condition is not true or the variable referenced by the condition is not null.

For example, suppose that the source data has four null values for Dose (indicated by empty cells in the following table):

PATIENT_ID	DRUG	Dose
A	Thiamine	
A	Niacin	15
B	Thiamine	
B	Niacin	25
C	Calcium	
D	Thiamine	15
E	Thiamine	25
F	Niacin	
G	Calcium	25

For a condition where Dose is greater than or equal to 20 and **Include Null values** is selected, the following patients are retrieved:

A

B

C

E

F

G

When you specify that values for a query should include nulls, and the condition is preceded by the NOT operator, the query retrieves patients with rows for which the condition is not true and the variable used in the condition is not null.

If you specify NOT 3 (meaning not condition 3, which is Dose is greater than or equal to 20), and you have selected **Include Null values**, the following patients are retrieved:

A

D

Reports tab

Reports page

Reports allow you to view source data in a table or graph. When you create a report, select the variables for the rows and columns of the report, and break down the variables at multiple levels, such as counts of patients for males and counts of patients for females within each age group. You can also indicate how to aggregate multiple values for variables in the report.

You can create various types of reports, from line listings to summary reports showing statistics for cross-tabulations of variables. If a report includes patient IDs or counts, you can use the drill down feature to view patient details.

The application creates a report output when you apply a report definition to the data in a patient group.

In the Empirica Healthcare Analysis application, various reports are generated based on a set of report definitions. The Reports tab has the following views:

- Report Definitions (default view)
- Report Outputs

Running a report as a background job

You have the following options for running a report:

- Run the report as a background job. For more information, see Creating a report output.
- Run the report while you wait. For more information, see Running a report.

When you select the **Create Output** option for a report definition, the report output is created as a background job. But when you click **Run**, the report output is created while you wait and as the results are displayed.

Note: Oracle recommends that you run reports as background jobs.

Report Definitions view

The Report Definitions view displays a list of all existing report definitions.

A report definition is the specification of the format of a report and any restrictions on the data that appears in the report. Users with appropriate permissions can create report definitions and make them available to other users.

A report definition specifies:

- The variables from the source data to appear as columns in the report along with their breakdown details based on the other variables.
- The variables to provide unique keys for the report. The report has one row for each unique key value or a combination of key values.
- The way to aggregate multiple values in individual cells.
- The subset of rows to appear in the report based on the conditions specified by a SQL WHERE clause.

You can create report definitions on an as-needed basis.

A report definition is saved with a unique identifying number and user-assigned descriptive information.

General activities

The following links and filters appear at the top of the page and affect the entire page.

- Create Definitions
- Columns
- Print
- Download
- Select Rows
- Filter by Project or Configuration

Row-specific activities

The following menu options are available from the row menu, located in the left-most column of the table, and affect an individual row in the table:

- Create Output
- Run
- Edit
- Rename
- Copy
- Publish
- Delete

Field descriptions—Report Definitions page

Field	Description
Category	Category of the report definition: • Ad-hoc • Standard
Configuration	Name of the configuration associated with the report definition.
Created	Date and time when the report definition was created.
Created By	User name of the person who created the report definition.
Definition	Name of the report definition.
Description	Description of the report definition.
ID	Identifier that was assigned to the report definition when it was created. A report definition ID is unique and is not reused if it is deleted.
Modified	Date and time when the report definition was last modified.
Modified By	User name of the person who modified the report definition.
Project	Project associated with the report definition.

Report Structure

A report consists of one row for each combination of variables that have been specified as row variables in the report definition. The remaining columns in the report are determined by the column variables specified in the report definition. You can create multiple levels of column variables.

Note: The name of the row variable is represented as a column in the report.

For example, in a report excerpt as shown below, there are:

- Two row variables—Age Group and Gender.
- Three column variables—Drug, Event, and Patient ID.

Age Group	Gender	Drug			
		Amoxicillin		Ampicillin	
		Event		Event	
		78701:Nausea with vomiting	78702:Nausea alone	78701:Nausea with vomiting	78702:Nausea alone
		Patient ID	Patient ID	Patient ID	Patient ID
		N (U)	N (U)	N (U)	N (U)
18-24	F	424	341	11	14
18-24	M	180	97	0	0
25-34	F	485	478	13	15
25-34	M	156	101	1	0
35-44	F	269	284	6	3
35-44	M	75	60	1	0

The column variable that determines the content of the cells of the report is an analysis variable. In this example, Patient ID is the analysis variable, and the content is defined as the unique count of patients that is represented as **N(U)**.

Row variables or column variables that group values for another variable are breakdown variables. In this example, Age Group, Gender, Drug, and Event are the breakdown variables for the Patient ID variable.

All variables in a report definition except the analysis variable are breakdown variables. The ALL column is not based on values of other columns in the report.

The Report Preview page shows data for only the first ten patients of the selected patient group or the first ten patients found by the query.

Creating a report definition manually

Creating a report definition

Creating a report definition consists of the following steps:

1. Selecting a patient group.
2. Naming the report definition.
3. Defining the rows and columns, and selecting the data source.
4. Defining breakdown details.
5. Defining content details.
6. Changing report attributes.
7. Changing report descriptors.
8. Saving the report.

Step 1—Selecting a patient group

Select a patient group before you run, edit, or copy a report definition. The report runs for the patients from the selected patient group.

When a patient group is selected, the report definitions based on data configurations compatible with the selected patient group are listed.

1. Select the **Reports** tab.

   The Report Definitions page appears.

2. Click **Browse** and select a patient group to which to apply the report definition.

3. Click **OK**.

   The Report Definitions page appears.

   Note: If the patient group for which you are viewing the report has any excluded patients (patients marked as Excluded in the Review Input section of the Patient Details page), then these excluded patients are not included in the report.

4. Continue to Step 2—Naming the report definition.

Step 2—Naming the report definition

Previous step	Step 1—Selecting a patient group.
Description	In this step, you provide a name and description of the report definition, along with the project associated with the report. Later in the wizard, you define rows and columns to include in the report, select the data source from which the data is retrieved for the selected rows or columns, and preview the output subset.

1. On the Report Definitions page, click **Create Definition.**

   The Create Definition page appears.

2. Fill in the fields, and click **Save.**

   The Edit Report Columns page appears.

3. Continue to Step 3—Defining the rows and columns, and selecting the data source.

Step 3—Defining the rows and columns and selecting the data source

The Edit Report Columns page appears when you create or edit a report definition, or when you click Report Columns on one of the report definition pages. On this page, you perform the following tasks:

1. Specify the variables and values (from the source data) to define rows and columns of the report.
2. Specify the labels of report columns.
3. Specify the aggregation method (such as count, percentage, mean, or actual value) to be used to show values in the report.

A report can include columns from multiple tables that consider Patient ID as the unique key. As a result, you can use only tables that contain a patient ID column in the reports.

Note: You can add an unlimited number of row variables, column variables, and hierarchical levels of column variables. For more information on the report format, see Report Definitions view.

Previous step	Step 2—Naming the report definition.
Description	In this step, you define rows and columns to include in the report, select the data source from where the data will be retrieved for the selected rows or columns, and preview the output subset. Later in the wizard, you define breakdown details for the rows and columns variables.

To define a first row variable:

1. In the Select column heading to review/edit section, click the yellow **{New Row Variable}** link**.**

   The lower half of the page displays information about the row variable.

2. In the {New Row Variable} section, click **[Select]**, located after **Data Source**.

   The Edit Data Source dialog box appears.

3. Select a variable from the list. To view only the variables in a specific table, from the **Table** drop-down list, select the table.

4. Click **OK**.

   The row heading is updated to match the variable you selected.

5. Do one of the following:

   - To define breakdown details, continue to Step 4—Defining breakdown details.
   - To define content details, continue to Step 5—Defining content details.

- To change report attributes, continue to Step 6—Changing report attributes.
- To change report descriptors, continue to Step 7—Changing report descriptors.
- To save the report, continue to Step 8—Saving the report.

To define a first column variable:

1. In the Select column heading to review/edit section, click the yellow **{New Column Variable}** link**.**

   The lower half of the page displays information about the column variable.

2. In the {New Column Variable} section, click **[Select]**, located after **Data Source**.

   The Edit Data Source dialog box appears.

3. Select a variable from the list. To view only the variables in a specific table, select the table from the **Table** drop-down list.

4. Click **OK**.

   The column heading is updated to match the variable you selected.

5. Do one of the following:

   - To define breakdown details, continue to Step 4—Defining breakdown details.
   - To define content details, continue to Step 5—Defining content details.
   - To change report attributes, continue to Step 6—Changing report attributes.
   - To change report descriptors, continue to Step 7—Changing report descriptors.
   - To save the report, continue to Step 8—Saving the report.

To add more row variables or column variables:

1. In the Select column heading to review/edit section, select a row or column variable.

   The selected row or column variable is highlighted in yellow.

2. Do one of the following:

   - To insert a row or column to the left of the highlighted row, click **Insert Left**.
   - To insert a row/column to the right of the highlighted row, click **Insert Right**.
   - To insert column above the highlighted column, click **Insert Above**.
   - To insert the column below the highlighted column, click **Insert Below.**

To edit a row variable or a column variable:

1. In the Select column heading to review/edit section, select a row/column variable.
2. Modify the data source and breakdown details as needed.

To preview the output of the report:

After you have finished defining a row or a column variable, you can preview the output of the report.

- Click **Preview**, located after the **Select column heading to review / edit section**.

Note: If any row or column variable appears in italicized red text, you cannot preview the report definition. This restriction occurs because you have not defined breakdown details or some other component correctly.

Step 4—Defining breakdown details

Breakdown details are the specification of a value, a range, or a set of values or ranges for a breakdown variable You can refine your data for a particular row variable or column variable based on the breakdown variables.

- If a breakdown variable is a text or date variable, define the breakdown details using individual selected values, distinct values, or grouped values.
- If a breakdown variable is a numeric variable, define breakdown details using cutpoints, which are the ranges defining the upper and lower limits of the values.

Note: If a column variable has ten or fewer values, all values are used by default for the breakdown variable. For a row variable, all values are used by default as distinct (unique) values for the breakdown variable. You can modify these default breakdowns as needed.

For best report performance, when possible, select the analysis variable from the same table as the row variables.

Previous step	Step 3—Defining the rows and columns, and selecting the data source.
Description	In this step, you define breakdown details for the row and column variables. Later in the wizard, you define content details for the columns variables.

1. On the Edit Report Columns page, click **[Select],** located after **Breakdown Details.**

   The Breakdown Details pop-up window appears.

2. Type the breakdown variables.

For text or date variable, you can:

- **Define breakdown details by individual values.**

  If the breakdown details are defined by individual values, the report includes a column or row for each value that you select. This option results in a static list of breakdown values. The list does not automatically adjust to changing source data. When **Individual Values** is selected, the default label for values appears as the exact values from the source data. You can modify the label that appears in the report. Editing the labels has no effect on the source data.

  i. In the Breakdown Details dialog box, select **Individual Values**.

  ii. To include a column or row to represent the missing values for the variable, select **Include a category for NULL values**.

  iii. To include a column or row labeled ALL in the report to represent all the values that are selected explicitly as breakdown values, select **Include a category for ALL selected values**.

For example, if breakdown values for drugs are DrugA and DrugB, the ALL column or row represents DrugA or DrugB.

Notes:

- The ALL category includes null values only if a NULL category is included. The ALL category does not include other values even if the OTHER category is included.
- For reports that display percentages, the report must include an ALL category.

iv. To include the OTHER category for values that are not selected explicitly as breakdown values, select **Include a category for all unselected values**. This option considers only terms that are in the group of patients against which the report is run.

For example, suppose that the breakdown variable is Generic Name. If you select the individual values DrugA and DrugB and then select this option, the OTHER category represents generic names that are not DrugA or DrugB, but that were taken by patients in the patient group.

The OTHER category does not include null values, regardless of whether a NULL category is included.

- **Define breakdown details by all distinct values**.

If the breakdown details are defined by all distinct values, the report includes a column or row for each distinct value of the variable. This option uses the values available at the time you run the report.

i. In the Breakdown Details dialog box, select **Every Distinct Value**.

ii. To include an ALL category in the report to represent all distinct values, select **Include a category for ALL selected values**.

For example, if all values include genders Female and Male, then the ALL category represents Female or Male.

Notes:

- The ALL category includes null values only if a NULL category is included.
- For reports that display percentages, the report must include an ALL category.

iii. To include a NULL category to represent the missing values for the variable, select **Include a category for Null values**.

- **Define breakdown details by grouped values**.

If the breakdown details are defined by the grouped values, the report includes a column or row for each group that you define. This option results in a static list of breakdown values that does not automatically adjust to changing source data.

i. In the Breakdown Details dialog box, select **Grouped Values**.

ii. To include a column or row to represent the missing values for the variable, select **Include a category for NULL values**.

iii. To include an ALL category in the report to represent all the values that are selected explicitly for the breakdown groups, select **Include a category for ALL selected values**.

For example, if breakdown groups include DrugA, DrugB, and DrugC, the ALL category represents DrugA, DrugB, or DrugC.

Notes:

- The ALL category includes null values only if a NULL category is included. It does not include other values even if an OTHER category is included.
- For reports that display percentages, the report must include an ALL category.

iv. To include an OTHER category for values that are not selected explicitly to be included in the breakdown groups, select **Include a category for all unselected values**. This option considers only terms that are in the group of patients against which the report is run.

For example, suppose that the breakdown variable is Generic Name. If you select the individual values DrugA and DrugB and then select this option, the OTHER category represents generic names that are not DrugA or DrugB, but that were taken by patients in the patient group.

The OTHER category does not include null values, regardless of whether a NULL category is included.

For a numeric variable, you can define breakdown detail by cutpoints:

1. In the Breakdown Details dialog box, type the values.
2. To include a column or row to represent null values for the variable, select **Include a category for Null values**.
3. To include a column or a row in the report to represent all values for the variable (including null values, only if you select **Include a category for Null values**), select **Include a category for ALL selected values**.
4. For reports that display percentages, the report must include a row or column for ALL.
5. After you have finished click **Save**.
6. Do one of the following:
   - To define content details, continue to Step 5—Defining content details.
   - To change report attributes, continue to Step 6—Changing report attributes.

- To change report descriptors, continue to Step 7—Changing report descriptors.
- To save the report, continue to Step 8—Saving the report.

Step 5—Defining content details

Content details specify the aggregation method for values. The aggregation method determines how the values of an analysis variable appear in a report. For example, you might want to show an actual value, a count of records with a value, a percentage of records with the value, and so on. You can specify multiple aggregation methods for an analysis variable; each aggregation method is represented in the report.

When you add an analysis variable to a report, Value appears by default as the aggregation method.

Value stands for the first value, which is the default aggregation method. (It is used when no other aggregation method is selected.) When you change the aggregation method, an abbreviation of that aggregation method appears in the table cell. For example, if you select Count as the aggregation method, N appears in the table cell.

Some aggregation methods compute counts and percentages. Counts and percentages for the ALL column or row are not computed by totaling the other columns or row. Therefore, if a patient is counted in multiple columns or rows, a count in the ALL column or row is not necessarily the same as the total of counts in other columns or rows of the report.

Previous step	Step 4—Defining breakdown details.
Description	In this step, define content details, such as Count, Row %, and others, for individual cells to specify how to display values of an analysis variable. Later in the wizard, you define report attributes to change drill down access for report cells, or restrict data using a SQL where clause.

1. On the Edit Report Columns page, from the table under **Select column heading to review/edit** section, click **Value** for a particular cell.

   The Content Details section appears at the bottom of the page.

2. Click **[Select]**, located after **Content Details**.

   The Content Details dialog box appears.

3. Select the aggregation methods.

   - **For text or date variables, select one or more of the following options:**

     The abbreviation appears in the report column heading.

Field	Abbreviation	Description
Count	N	Number of non-NULL values.
Count (Unique)	N (U)	Number of unique non-NULL values.

Row %	Row %	Count of non-unique values for one of the following: • The rows and columns. • A row for all columns. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL column.
Column %	Column %	Count of non-unique values for one of the following: • A column and a row. • A column for all rows. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL row.
Overall %	Overall %	Count of non-unique values for one of the following: • A column and a row. • All columns and all rows. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL column and row.
Row % (Unique)	Row % (U)	Count of unique values for one of the following: • A row and a column. • A row for all columns. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL column.
Column % (Unique)	Col % (U)	Count of unique values for one of the following: • A column and a row. • A column for all rows. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL row.
Overall % (Unique)	% (U)	Count of unique values for one of the following: • A column and a row. • All columns and all rows. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL column and row.

All values	ALL	All values, including null values, separated by commas. NULL values are shown as blanks, so you might see a list of values such as **DrugA, , DrugC**.
All values (excluding NULLs)	All	All values except NULL values, separated by commas.
Unique values	ALL (U)	Unique values in ascending order, including NULL values, separated by commas. NULL values are represented in the list by the string **NULL**.
Unique values (excluding NULL)	All (U)	Unique values in ascending order, excluding NULL values, separated by commas.
Unique values and counts	ALL (N)	Unique values in ascending order, including NULL values, separated by commas. The count of the value appears after each value. NULL values are represented in the list by the string **NULL**.
Unique values and counts (Excluding NULL)	All (N)	Unique values in ascending order, excluding NULL values, separated by commas. The count of the value appears after each value.
First value	Value	First non-NULL value from the database, for the patient group or query being used. If you do not specify an aggregation method, this method is used by default.

- **For numeric variables, select one or more of the following options:**

  For an analysis variable with a data type of NUMBER, select one or more of the following options. The abbreviation appears in the report column heading.

Field	Abbreviation	Description
Count	N	The number of non-NULL values.
Count (Unique)	N (U)	The number of unique non-NULL values.
Row %	Row %	Count of non-unique values for one of the following: • The rows and columns. • A row for all columns. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL column.

Column %	Col %	Count of non-unique values for one of the following: • A column and a row. • A column for all rows. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL row.
Overall %	%	Count of non-unique values for one of the following: • A column and a row. • All columns and all rows. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL column and row.
Row % (Unique)	Row % (U)	Count of unique values for one of the following: • A row and a column. • A row for all columns. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL column.
Column % (Unique)	Col % (U)	Count of unique values for one of the following: • A column and a row. • A column for all rows. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL row.
Overall % (Unique)	% (U)	Count of unique values for one of the following: • A column and a row. • All columns and all rows. In the Breakdown details section, **Include a category for ALL selected values** must be selected for ALL column and row.
Sum	Sum	Sum of all the values.
Mean	Mean	Mean of all the values.
Median	Median	Median of all the values.
Standard Deviation	SD	Standard deviation of all the values.

1st Quartile	Q1	First quartile of all the values.
3rd Quartile	Q3	Third quartile of all the values.
Min	Min	Minimum value.
Max	Max	Maximum value.

4. After you have finished click **OK**.
5. Do one of the following:
   - To change report attributes, continue to Step 6—Changing report attributes.
   - To change report descriptors, continue to Step 7—Changing report descriptors.
   - To save the report, continue to Step 8—Saving the report.

Step 6—Changing report attributes

Report attributes allow you to:

- Drill down the report column variables that show counts of patients, or on elements of graphs based on such reports.
- Restrict a report to data that meets a specified condition by using a SQL Where clause.

  For example, to create a report that includes males who reside in the United States and are 18 to 55 years old. A report restriction is applied, if specified, after the report is generated and before the report is displayed.

Previous step One of the following steps:

- Step 3—Defining the rows and columns, and selecting the data source.
- Step 4—Defining breakdown details.
- Step 5—Defining content details.

Description This step is optional.

In this step, you edit the drill down access for the report cells, and restrict the data using a SQL WHERE clause.

Later in the wizard, you edit the report descriptors.

---

1. On the Edit Report Columns page, select **Report Attributes**, located in the row of radio buttons at the top of the page.

   The Edit Report Attributes page appears.

2. To edit the drill down access for the report cells:

   A. Click **Generate Drilldown Information**.

   The Generate Drilldown Info dialog box appears.

   B. Select **Yes** or **No**, and click **OK.**

3. To define a criteria to restrict the report attributes to meet your requirements:

   A. Click **Restrict by SQL WHERE Clause**.

   The Restrict by SQL WHERE Clause page appears.

   B. To view the available columns in the report definition, click **Show columns**.

   The Select Table Columns dialog box appears.

   C. Type the SQL query in the SQL WHERE Clause field, and click **Apply**.

4. Do one of the following:

   - To change the report descriptors, continue to Step 7—Changing report descriptors.

- To save the report, continue to Step 8—Saving the report.

Step 7—Changing report descriptors

Previous step	One of the following steps: • Step 3—Defining the rows and columns, and selecting the data source. • Step 4—Defining breakdown details. • Step 5—Defining content details. • Step 6—Changing report attributes.
Description	This step is optional. In this step, you edit name, description, and category assignment for the report, and view the XML of the report definition that you are creating through this wizard. Later in the wizard, you save the report.

1. On the Edit Report Columns page or the Edit Report Attributes page, click **Report Descriptors**, located in the row of radio buttons at the top of the page.

   The Edit Report Descriptors page appears.

2. Fill in the fields, and click **Save**.

3. Continue to Step 8—Saving the report.

Field Description—Edit Report Descriptors page

Field	Description
Name of Report	Name of the report definition.
Description of Report	Description of the report definition.
Category	Category of the report definition: • Ad-hoc • Standard
Project	Project to which to assign the report definition.
Status	Status of the report definition: • **Ready—**The report definition is valid and can be run. A report definition is considered valid if it has at least one row variable, at least one column variable, and no error messages for the report variables. • **Under Development**—The report definition is not ready to be run.
Configuration	Description of the source data of the report definition.

Step 8—Saving the report

Previous step	One of the following steps • Step 3—Defining the rows and columns, and selecting the data source. • Step 4—Defining breakdown details. • Step 5—Defining content details. • Step 6—Changing report attributes. • Step 7—Changing report descriptors.
Description	In this step, you save the report definition.

- On the Edit Report Columns page, save the report by performing one of the following actions:
  - **Save**—Save the report definition without running it. If you are working on a complex report definition, you might want to click **Save** periodically. The report definition remains visible so that you can continue working on it.

    Note: If you are editing a report definition that you did not create, the **Save** button is not available.
  - **Save As**—Save the report definition with a different name, without running it. The report definition remains visible so that you can continue working on it.
  - **Save & Run**—For a valid report definition, save the report definition and run it. A report is considered valid if it includes at least one row variable and one column variable, and the report definition has no error messages.

Creating a report definition from XML

Obtaining the XML for a report definition

1. Select the **Reports** tab.
2. Select a patient group.
3. Select the row menu () for a report definition, and click **Edit**.

   The Edit Report Columns page appears.
4. Select the **Report Descriptors** radio button.
5. Copy the text in the **XML** field. You paste the text into the XML Definition field when you create a report definition from XML.

Creating a report definition from XML

A report definition is stored as XML (eXtensible Markup Language). You can send the XML representation of the report to other users, who can create new reports using the XML.

Prerequisites

- You must be a superuser.
- Obtain the XML for a report definition.

To create a report definition from XML:

1. Select the **Reports** tab.

   The Report Definitions page appears.

2. If a patient group is not selected:

   A. Click **Browse** and select a patient group to which to apply the report definition.

   B. Click **OK**.

   Note: If the patient group for which you are viewing the report has any excluded patients (patients marked as Excluded in the Review Input section of the Patient Details page), the excluded patients are not included in the report.

3. Click **Create Definition from XML**.

   The Create Definition from XML page appears.

4. Fill in the fields according to the following table.

   The values that you specify appear on the Reports tab, where you can sort them. For example, you can group reports into categories and then sort by category on the Reports tab so that all reports of that category are listed together.

5. In the **XML Definition** field, paste the copied XML, and click **Save**.

   The application creates the report definition, which appears on the Report Definitions page.

Field descriptions—Create Definition from XML page

Field	Description
Name	Name of the report.
Description	Description of the report.
Category	Category containing the report.

Creator	Read-only. Name of the user who created the report.
Data	Read-only. Description of the source data.
Status	Read-only. **Under Development** appears, indicating that the report is not ready to be run until you save it.

Creating a report output

When a report definition is applied to a patient group to create a report, the report output is saved in the Report Outputs view.

1. Click the **Reports** tab.
2. Select the row menu (), and click **Create Output**.

   The Create Output page appears.
3. Fill in the fields, and click **Save.**

   The report output appears in the Report Outputs view with a status of Running. After the report runs successfully, the status change to Completed.

Running a report and saving the output

You can run a report only when a patient group is selected.

1. Click the **Reports** tab.

2. Select the row menu (), and select **Run**.

   The Running Report page appears with the status of the running report:

   - **Working**—The report is running. Click **Cancel** to cancel to the report.
   - **Failed**—The running of the report failed. Click **Continue** to return to the Report Definitions page.
   - **Cancelled—**The report is cancelled. Click **Continue** to return to the Report Definitions page.

   After the report finishes running, the Display Report page appears with report results.

   If an out-of-memory error occurs while running the report, turn off the drill down in the report definition and try to run the report again. When the drill down option is off, less memory is used to run the report; however, the report may take longer to run.

   For more information on using drill down option, see Step 6—Changing report attributes.

3. Click **Save Output**.

4. Fill in the fields, and click **Save**.

   The report output is saved and appears in the Report Outputs view.

Report Outputs view

A report output is the result of applying a report definition to the patients that are in a patient group, and saving the result. A report output is static. If a change occurs to a report definition or to the set of patients against which the report was run to produce the output, you might want to run the report again and save another report output.

When you apply a report definition to the data in a patient group, a report output is created. You can view the report output in tabular or graphical format and save it for later review.

The Report Outputs view displays a list of report definitions that are run.

General activities

The following links and filters appear at the top of the page and affect the entire page.

- Columns
- Print
- Download
- Select Rows
- Filter by Project or Configuration

Row-specific activities

The following menu options are available from the row menu, located in the left-most column of the table, and affect an individual row in the table:

- View
- Rename
- Publish
- Delete

Field descriptions—Report Outputs page

Field	Description
Category	Category of the report output: • Ad-hoc • Standard
Configuration	Name of the configuration associated with the report definition.
Created	Date and time when the report definition was created.
Created By	User name of the person who created the report definition.
Definition	Name of the report definition that is used to create the report output.
Description	Description of the report definition.

ID	Identifier that was assigned to the report output when it was created.
Output	Name of the report output.
Patient Group	Name of the patient group associated with the report definition.
Project	Name of the project associated with the report definition.
Status	Status of the report output: • **Completed—**The report completed successfully. • **Running**—The report is running. • **Error Occurred —**A job failed, resulting in a failed report output. • **Cancelled—**The creation of the report output was cancelled.

Viewing a report graph

Viewing a report graph

You can view a report output in a graphical representation. In the Empirica Healthcare Analysis application, different types of graphs are available depending on the types of columns in the report.

Graph type	Report variables
Aggregate bar graph	All column variables or all column variables except one are numeric. The patient ID does not need to be in the report. Use link: **Bar graph (where rows are aggregate values)**
Detail bar graph	The row variable in the report is the patient ID. Use link: **Bar graph (where rows are detail records)**
Box plot graph	The row variable in the report is the patient ID, and at least one column variable is numeric. Use link: **Box plot (where rows are detail records)**
Scatter plot graph	The row variable in the report is the patient ID, and at least two column variables are numeric. Use link: **Scatter plot (where rows are detail records)**

You can set the graph key and color palette based on the analysis results. It does not affect the report graphs. The report graphs appear in the color unless you choose the display option **Use gray-scale instead of colors**.

Note: A numeric variable is one for which source data is numeric or for which content detail is specified as a numeric value, such as count or percentage.

To view a report graph:

1. Click the **Reports** tab.
2. Click the **Report Outputs** link located under the Reports tab.

   The Report Output page appears.
3. Select the row menu (), and select **View**.

   The Display Output page appears
4. Click **Choose Graph**.

   The Choose Graph page appears.
5. Select the type of graph from the Available graphs list, and click **Display**.

Selecting the type of graph

Selecting an aggregate bar graph

In an aggregate bar graph, the row variable of the report is represented on the y-axis, and each numeric column variable in the report is represented on the x-axis. A key appears below the graph describing each bar.

The display options that you set are used for only the current display of the graph.

To select an aggregate bar graph:

1. On the Choose Graph page, select **Bar graph (where rows are aggregate values)**.

   The Aggregate Bar Graph page appears.

2. Type **X** and **Y** labels to specify the information that is represented by the axes.

3. To display the patient counts at the ends of bars, select **Show counts at the ends of bars**.

   If the report shows numeric values that are not patient counts and you select this option, the values are shown instead of counts.

4. To swap rows and columns of the report, select **Transpose rows and columns**.

   You can also set any other display options.

5. Click **Display**.

   The graph appears.

6. To drill down to the patients of a bar, select a bar that represents patient counts.

Note: If multiple graphs appear on the same page, do not click any graph until all the graphs appear.

Selecting a detail bar graph

The detail bar graph determines the row variables and column variables of the report that are represented by the y-axis, and the subject counts that are represented on the x-axis.

A subject count appears for each primary variable that you select as a display option. You can break down the patient counts for the primary variable by specifying a secondary variable. If you specify a secondary variable, a graph key indicates the meaning of each bar.

You can also create multiple graphs, one for each value of a row or column variable that you specify as a subset variable.

To select a detail bar graph:

1. On the Choose Graph page, select **Bar graph (where rows are detail records)**.
2. Select a **subset variable**.

   One graph for each value of the selected variable will be generated.
3. Select a **primary variable**, to be represented on the x-axis.
4. You may select a **secondary variable**. Each bar in the graph represents a value of the secondary variable, as identified by the color key below the graph.
5. To set the display options, fill in the fields.

Field	Description
Labels X and Y	Label the x-axis, and y-axis to clarify what is represented by the axes.
Show counts at the ends of bars	Show patient counts at the end of bars in the graph.

Note: The display options are used only for the current display of the graph.

6. Click **Display.**
7. To view the exact values represented by a bar in the graph, point to the bar.
8. To drill down to the patients for a bar, select the bar.

Note: If multiple graphs appear on the same page, do not click any graph until all the graphs appear.

Selecting a box plot graph

A box plot, also known as a box-and-whisker plot, is the plotting of data points against horizontal and vertical axes to show the distribution of a continuous variable.

In a box plot:

- The box represents the middle 50 percent (approximately) of the numeric values.
- A horizontal line within the box represents the median of all values
- The top end of the box represents the upper quartile, which is the median of the ordered set of values that are greater than the overall median.
- The bottom end of the box represents the lower quartile, which is the median of the ordered set of values that are less than the overall median.

The interquartile range, which is the difference between the upper quartile and the lower quartile, is a measure of the spread of the distribution. The relative distances of the upper and lower quartiles from the median describe the shape of the distribution of data. In the box plot:

- The whisker above the box plot extends from the upper quartile to the highest actual value that is within the (75th percentile + 1.5 * (interquartile range)).
- The whisker below the box plot extends from the lower quartile to the lowest actual value that is within the (25th percentile - 1.5 * (interquartile range)).
- Outliers are plotted as individual points on the graph. An outlier is considered to be a value that falls outside of the upper or lower whisker.

Note: Points in a box plot are displayed at small random offsets from the center line in a process known as jittering. This display ensures that if two records have the same value, a point is likely to be displayed for each of them.

Several box plots might appear in a single graph if you select a secondary variable. In that case, a key appears below the box plot to relate the individual box plots to the values of the secondary variable.

To select a box plot:

1. On the Choose Graph page, select **Box plot (where rows are detail records)**.
2. You may select a **subset variable**.

   One graph for each value of the selected variable is generated.
3. Select a **primary variable** from a list of numeric variables in the report definition. Each point in the box plot represents a value of the primary variable.
4. You may select a **secondary variable**. Each box represents a value of the secondary variable, as identified by the color key below the graph.
5. To set the display options, fill in the fields.

Field	Description
Labels X and Y	Label the x-axis and y-axis.
Show all points as overlay	Show all points on the box plot. If cleared, only points outside the upper and lower whiskers will be shown. *Note*: Points in a box plot are jittered, so that if two records have the same value, a point is likely to be displayed for each of them.

Note: The display options are used only for the current display of the graph.

6. Click **Display.**

7. If you point to a region of the graph, the following information appears:

   - The region of the box (Upper Outlier, Upper Whisker, Upper Box, Lower Box, Lower Whisker, or Lower Outlier).
   - The value of the secondary variable, if any.
   - The count of data points for the primary variable (and secondary variable, if any) for that region of the box.

   The count may be more than the number of patients as there may be more than one data point for the same patient. For example, a patient may have multiple adverse events, where each adverse event is a separate data point.

8. To drill down to the patients of that region, click a region—Upper Outlier, Upper Whisker, Upper Box, Lower Box, Lower Whisker, or Lower Outlier, of the graph.

Note: If multiple graphs appear on the same page, do not click any graph until all the graphs appear.

9. If a patient has a value on the boundary between any regions, not including the Upper Outlier or Lower Outlier regions, the patient ID is included when you drill down on either of the regions. Patient with values on the boundary of the Upper Whisker or Lower Whisker are not included when you drill down on the Upper Outlier or Lower Outlier regions.

Selecting a scatter plot graph

A scatter plot is the plotting of data points against horizontal and vertical axes to show the correlation between two or more continuous variables. You select a variable to plot against the x-axis of the scatter plot, and one or two variables to plot against the y-axis.

You can select a variable by which to subset data. For example, you might want to view a different scatter plot for each diagnosis.

You can specify an inner breakdown variable if you want to use a color key to associate each dot on the graph with a particular value. For example, if you want to know the particular patient ID associated with each dot, you can use the patient ID as the inner breakdown variable.

Notes:

- A data point is not plotted for a record if either the X variable or the Y variable is empty.
- If you select two y-axis variables, each of them uses a different shape, as shown in the color key below the scatter plot.
- If the X and Y values for one record coincide with the X and Y values for another record, the two data points are plotted in the same position. For such cases, you cannot distinguish whether that data point represents one record or several records.

The color key below the scatter plot shows which value of the inner breakdown variable each dot represents.

To select a scatter plot:

1. On the Choose Graph page, click **Scatter plot**.
2. You may select a **subset variable**.

   One scatter plot for each value of the subset variable will be generated.
3. Select an **X** variable, select a numeric variable to be plotted against the x-axis.
4. Select a **Y** variable, select a numeric variable to be plotted against the y-axis.
5. You may select another **Y** variable. Each of the two Y variables is plotted against the y-axis. Each Y variable corresponds to a different-shaped point, as shown in the key below the scatter plot.
6. You may select a text variable as an inner breakdown variable. The color key associates each point on the scatter plot with a particular value of the inner breakdown variable.
7. To set the display options, fill in the fields.

Field	Description
Labels X and Y	Label the x-axis and y-axis to clarify what is represented by the axes.
Include linear regression lines	Include a linear regression line to help identify trends in the data.
Include 45-degree line and matched axes	Include a 45-degree line in the scatter plot. The line is useful if you are plotting two similar variables and want to distinguish between the two.

8. To drill down to patient details, click within the graph and draw a red rectangle around the data points for which you want to drill down. When you release the mouse button, a menu appears, and you can drill down. A single point in the graph might represent several data points.

Note: If multiple graphs appear on the same page, do not click any graph until all the graphs appear.

Single- and multi-patient timelines

Single-patient timelines

A single-patient timeline is a visual presentation of a patient's drug exposures, diagnoses, and procedures. A single-patient timeline shows source data and does not use the temporal parameters from analysis runs. You can open single-patient timelines from patient counts, which appear in many places in the application; the context from which you open the timeline determines the data configuration that is used. For example, if you open a single-patient timeline from analysis results, the current configuration is the configuration that was used for the analysis.

The graph consists of the following axes:

- **Horizontal axis**

  Tick marks for dates appear on the horizontal axis. The start and end of the horizontal axis vary by patient. The axis applies to all drugs, diagnoses, and procedures.

- **Vertical axis**

  Values for the types of data that you choose to view appear on the vertical axis. Each value appears along a gray horizontal line in the graph.

  Long values that do not fit in the allotted space include an ellipsis (...) in the middle of the value.

Note: The **Display Patient Timelines as applet** preference determines whether the presentation is an interactive display or a static JPEG image. Tooltips appear only if the graph is an interactive display.

Data points and symbols

A prescription for a drug is represented as a duration, which appears as a horizontal line with small vertical segments at each end. If prescriptions for the same drug overlap, you can configure the graph to stagger the symbols for the prescriptions so that the prescriptions can be differentiated. For information about configuring the graph, see Viewing and configuring single-patient timelines.

Each occurrence of a diagnosis or procedure is represented as a data point, which appears as one of a variety of symbols, as described in the following table.

Drugs

Symbols in this section appear in the key when the graph is configured to show drugs.

Symbol	Description
	A drug prescription.

Diagnoses

Diagnosis symbols appear in the key when the graph is configured to show diagnoses.

Symbol	Description
△	An occurrence of a diagnosis. If the configuration distinguishes inpatient from outpatient diagnoses, the symbol indicates an outpatient diagnosis.
□	(Appears if the configuration distinguishes inpatient from outpatient diagnoses.) An occurrence of an inpatient diagnosis. If the same diagnosis occurs on the same date as both inpatient and outpatient, only the inpatient symbol appears.

Procedures

Procedural symbols appear in the key when the graph is configured to show procedures.

Symbols in this section appear in the graph when the data configuration supports procedures.

Symbol	Description
△	An occurrence of a procedure. If the configuration distinguishes inpatient procedures from outpatient procedures, the symbol indicates an outpatient procedure.
□	(Appears if the configuration distinguishes inpatient procedures from outpatient procedures.) An occurrence of an inpatient procedure. If the same procedure occurs on the same date as both inpatient and outpatient, only the inpatient symbol appears.

Enrollment periods

If the data configuration is set up to include prescriptions starting outside of a patient's enrollment period and diagnoses or procedures occurring outside of a patient's enrollment period, this data appears in single-patient timelines.

Highlighting

Single-patient timelines support automatic highlighting and user highlighting. If you open the timeline from analysis results, terms that are used as the index event or outcome event in the analysis are automatically highlighted if the timeline is configured to show the same variable as was used for the index event or outcome event. For example, suppose that the index event was based on the generic name Aspirin. If you configure the graph to show generic names on the vertical axis, Aspirin is highlighted in the graph.

Automatic highlighting appears in a single-patient timeline only when you open the timeline from most analysis results. Automatic highlighted does not appear for excluded patients.

Highlighting color	Meaning	Notes
Yellow	Term of interest	(Appears in the key only if you open the timeline from descriptive analysis or evaluative analysis results.) Automatic highlighting shows the following information: • **For Risk-Outcome results**—A line with this highlighting indicates a value specified for the outcome event or index event. • **For Drug Utilization or Outcome Characterization results**—A line with this highlighting indicates a value specified for the outcome event. • **For evaluative results**—A line with this highlighting indicates a value specified for the outcome event or the index event. When you specify a procedure code while creating an analysis run, the procedure code is also highlighted.
Blue	User highlighting	(Always appears in the key.) You add user highlighting to the graph.

Note: You can add and remove user highlighting but not automatic highlighting.

Custom terms

When you open a single-patient timeline from an analysis run, the graph might include custom terms. A custom term is a name that you provide for a group of drugs, diagnoses, or procedures that you select for an index event or outcome event when defining the analysis run.

If the graphs shows custom terms, you can determine the specific term that a symbol represents by viewing the tooltip.

Additional information

Single- and multi-patient timelines use functionality from the Stottler Henke DataMontage visualization software, a third-party application. For more information, see the Stottler Henke documentation.

Multi-patient timelines

Unlike a single-patient timeline, which shows source data, a multi-patient timeline shows data from descriptive analysis and evaluative analysis runs. A multi-patient timeline shows data for multiple patients in a single graph. The graph includes the following information:

- Drugs, diagnoses, or procedures that were used as an index event or outcome event in the analysis.

  Any occurrences of the specified terms appear, not just the specific occurrence that is used to establish the index date or incidence date.

  For example, suppose that Patient A has the term specified in the run as the index event, but the occurrence of that term does not meet the lead-in criteria. In analysis results, the count of patients with an index date does not include Patient A. However, Patient A might be included in the initial count of patients. If you look at multi-patient timelines for patients included in the initial count, the graph includes Patient A and shows the index term.
- Enrollment start and end dates, as well as the drug eras and follow-up period generated by the analysis.

The graph consists of the following axes:

- **Horizontal axis**

  Tick marks showing the number of days relative to the alignment option that is used appear on the horizontal axis. The alignment option determines Day 0 in the graph. Data in the graph is placed in a position that is relative to Day 0.

  The horizontal axis encompasses all the enrollment start dates, enrollment end dates, and index terms or outcome terms for all patients on the current page of the graph.
- **Vertical axis**

  Individual patients appear on the vertical axis. A configuration option allows you to show or hide the patient IDs. For more information, see Viewing and configuring multi-patient timelines.

Note: The **Display Patient Timelines as applet** preference determines whether the presentation is an interactive display or a static JPEG image.

Data points and symbols

The following symbols might appear in the graph key, depending on the analysis type.

Symbol	Description
	Period of time that the analysis determined to be follow-up for the patient.
	Start of the patient's enrollment period.
	End of the patient's enrollment period.
	Diagnosis specified as the index event for the run.

	Diagnosis specified as the outcome event for the run.
	Procedure specified as the index event for the run.
	Procedures specified as the outcome event for the run.
	Prescription for the drug specified as the index event for the run.
	Prescription for drug specified as the outcome event for the run.
	Drug era for the drug that was specified as the index event for the run. **Note:** An analysis does not generate drug eras for prescriptions that start outside of a patient's enrollment.
	Drug era for the drug that was specified as the outcome event for the run.

Enrollment periods

The graph shows diagnoses and procedures that occur inside or outside a patient's enrollment period.

The graph shows all prescriptions, even if they start outside of the patient's enrollment period. The graph does not show eras for prescriptions that start outside of enrollment. Additionally, the symbol for a drug era ends at the earliest of the analysis period or enrollment period.

Evaluative analysis results

Because the graph shows terms that occur outside of enrollment, the following situations might occur in the results for an evaluative analysis run:

- A patient has the primary index term within enrollment and the comparator index term outside of enrollment. In this case, only the primary index term appears in the graph.
- A patient has the primary index term outside of enrollment and the comparator index term within enrollment. In this case, only the comparator index term appears in the graph.

Additional information

Single- and multi-patient timelines use functionality from the Stottler Henke DataMontage visualization software, a third-party application. For more information, see the Stottler Henke documentation.

Viewing and configuring single-patient timelines

For more information about single-patient timelines, see Single-patient timelines.

Prerequisites

- You must have the View Raw Safety Database permission.
- To use the interactive features of single-patient timelines, **Display Patient Timelines as applet** preference must be selected.

To view a single-patient timeline:

Note: Right-click options on the graph are available only if the graph is an interactive display, not if it is a static JPEG image.

---

1. Select the **Descriptive Analysis** or **Evaluative Analysis** tab.
2. Select the row menu () for an analysis run, and select **View Results**.

   Results of the analysis run appear. Several columns include patient counts with links.
3. Click a link in one of the columns with patient counts, and select **Single-patient Timelines** from the drop-down list.

   The Single-patient Timelines dialog box appears. It might take several moments for the graphics to appear the first time you use this feature after logging on.
4. To configure the appearance of the graph:

   A. Click **Configure**, located above the graph.

   The Single-patient Timelines Preferences dialog box appears.

   B. Fill in the fields according to the following table, and click **OK**.
5. To view a key for the graph, perform one of the following steps:

   - (In applet mode) Right-click a location in the graph, and select **Show Graph Key**.
   - (In static mode) Click **Show Key**, located at the top of the page.
6. (Available in applet mode only) To view a tooltip for a symbol in the graph, point to the symbol.
7. To select a different patient, select the patient from the **Current Patient** drop-down list.
8. To navigate through patients (if you selected a patient count with more than one patient):

A. Click the **Prior** and **Next** buttons at the top of the window to navigate through patients in the order of the list of patients.

B. If the patient count you selected had more than 2,000 patients, the drop-down list contains, as needed, **Previous Batch** and **Next Batch**. Select either option to navigate to another group of patients.

9. (Available in applet mode only) To find a value:

A. Right-click a location in the graph, and select **Find**.

B. Type a value, and click **OK**.

Only the part of values that are displayed on the vertical axis are searched.

Values that contain the string are highlighted in blue as horizontal lines in all graphs in the current set. Additionally, a vertical reference line is added to the first chronological occurrence of the text string.

If no matching values exist in the current graph, no highlighting is applied, even in other graphs for which the value exists.

Terms remain highlighted until you remove the highlighting.

C. To remove the highlighting, right-click the graph, and select **Remove User Highlighting**.

10. (Available in applet mode only) To highlight a line:

A. Right-click a line, and select **Highlight Term**.

The line is highlighted in blue in the current graph. Additionally, lines representing the same value (for a drug, diagnosis, or procedure) are highlighted in other graphs in the current set.

B. To remove the highlighting from all graphs, right-click the graph, and select **Remove User Highlighting**.

11. (Available in applet mode only) To view only data within the time period of a specific term:

A. Make sure that the graph has a reference line. The reference line represents the start of the time period. Click anywhere in the graph to create a reference line.

B. Right-click the place in the graph that you want to be the end of the time period, and select **Show Terms with Data in Region**.

The graph shows only data within the time period that starts with the reference line and ends with the spot you right-clicked. Diagnoses and procedures within the time period appear. If a prescription has any days within the time period, the entire prescription appears.

The display of only the current graph is affected.

C. To remove the effect of this option, right-click the graph and select **Show All Data**.

12. To print the graph, click **Print**, located above the graph.

If the graph is displayed as a static image, only one page of the graph display is printed.

Option descriptions—Single-patient Timelines dialog box

Option	Description
Layout	**Note:** Your selections in this section are saved for your user name for future login sessions across all data configurations.
Graph view for Drugs and Diagnoses	**Separate panels**—When selected, drugs and diagnoses appear in separate sections of the graph. **One panel**—When selected, drugs and diagnoses are intermingled in the same section of the graph.
Space between lines in graph	When **Condensed** is selected, less space appears between the lines in the graph.
Stagger overlapping intervals	When selected, the lines representing prescriptions are staggered so that you can differentiate between the components of overlapping prescriptions.
Content	For each type of data in the graph, a pair of fields appear: • One field determines the variable's values that appear on the vertical axis. • One field determine the variable's values that appear as tooltips. For example, you could configure the graph to show generic names on the vertical axis and show product names as tooltips. These fields include a <Same as run> option that refers to the variable that was used in the analysis run. **Note:** Your selections in this section are saved for your user name for future login sessions in only the current data configuration.

Note: If you open a single-patient timeline from the results of an analysis run, the variable drop-down lists might include a <Same as run> option. This option is available for each type of variable that was used to define the index event or outcome event in the analysis run.

Viewing and configuring multi-patient timelines

A graph appears when a patient count contains fewer than 10,000 patients.

For more information about multi-patient timelines, see Multi-patient timelines.

Prerequisites

- You must have the View Raw Safety Database permission.
- To use the interactive features of multi-patient timelines, **Display Patient Timelines as applet** preference must be selected.

To view a multi-patient timeline:

Note: Right-click options on the graph are available only if the graph is an interactive display, not if it is a static JPEG image.

---

1. Select the **Descriptive Analysis** or **Evaluative Analysis** tab.
2. Select the row menu () for an analysis run, and select **View Results**.

   Results of the run appear. Several columns include patient counts with links.
3. Click a link in one of the columns with patient counts, and select **Multi-patient Timelines** from the drop-down list.

   The Multi-patient Timelines dialog box appears. It might take several moments for the graphics to appear.
4. To configure the appearance of the graph:

   A. Click **Configure**, located above the graph.

   The Multi-patient Timelines Preferences dialog box appears.

   B. Fill in the fields according to the following table, and click **OK**.
5. To view a key for the graph, perform one of the following steps:

   - (In applet mode) Right-click a location in the graph, and select **Show Graph Key**.
   - (In static mode) Click **Show Key**, located at the top of the page.
6. (Available in applet mode only) To view a tooltip for a symbol in the graph, point to the symbol.
7. To navigate through sets of patients (if you selected a patient count that exceeds the allowed number of patients who can appear):

   - Click the **Prev** and **Next** buttons at the top of the window.

   or

Select a set from the **Current Set** drop-down list.

8. To print the graph, click **Print**, located above the graph.

   If the graph is displayed as a static image, only one page of the graph display is printed.

Option descriptions—Multi-patient Timelines dialog box

When you configure options, your changes are saved from session to session according to the following rules:

- For Risk-Outcome results—Your selections apply to all Risk-Outcome results that you view.
- For Drug Utilization and Outcome Characterization results—Your selections apply to all Drug Utilization and Outcome Characterization results that you view. For example, your selections for Drug Utilization results apply to a graph based on Outcome Characterization results.
- For evaluative analysis results—Your selections apply to all evaluative analyses that you view.

Option	Description
General	
Align by	Method for determining Day 0 for each patient in the graph. Other data in the graph for the patient is relative to that Day 0. For more information, see Alignment of multi-patient timelines.
Sort by	Order in which patients are listed on the vertical axis. Patients are listed in ascending order of the sorting option. For more information, see Sorting of multi-patient timelines.
Graph size	Amount of space between lines representing each patient's timeline.
Number of patients per page	Maximum number of patients that appear. The value must be between 1 and 500. If you view a multi-patient timeline for more patients than can be displayed on one page of the graph, patients are grouped into sets. Patients missing the alignment or sorting term are in their own set or in multiple sets and are last.
Show patient IDs	When selected, a patient ID appears along the vertical axis.
Drug Options	The following options appear even if the graph does not show drugs.
Stagger overlapping intervals	This option applies only to drugs and does not affect diagnoses or procedures. When selected, lines representing overlapping prescriptions for the

	index drug and outcome drug are separated so that you can differentiate among the prescriptions. For examples, see Examples of staggered prescriptions in a multi-patient timeline.
Show	This option applies only to drugs and does not affect diagnoses or procedures. The following information appears in the graph: • **Prescriptions**—Prescriptions use the prescription start and days supplied. Prescriptions appear even if drug eras were not created from them. • **Eras**—Drug eras generated by the analysis run appear. • **Prescriptions and eras**—Both prescriptions and drug eras appear. **Note:** An analysis does not generate drug eras for prescriptions that start outside of a patient's enrollment.

Alignment of multi-patient timelines

In the Multi-patient Timelines dialog box, you can specify the alignment of a multi-patient timeline. For more information, see Viewing and configuring multi-patient timelines.

Description

In the **Align by** option, you specify the method for determining Day 0 for each patient in the graph. Other data in the graph for the patient is relative to that Day 0.

Day 0 typically is visible on the horizontal axis. However, if you align by analysis start or absolute time, Day 0 might not be visible on the horizontal axis. The horizontal axis is wide enough to show data for the patient, but the analysis start and absolute time are not data for the patient.

Alignment by index term or outcome term does not consider whether the term is within the analysis or patient enrollment periods. For example, suppose that the index event is specified as Aspirin and Patient A's only prescription for Aspirin starts before the analysis period and before the patient enrollment period. If aligned by index term, the graph treats the start of that prescription as Day 0.

Alignment options

The following table options are available for alignment.

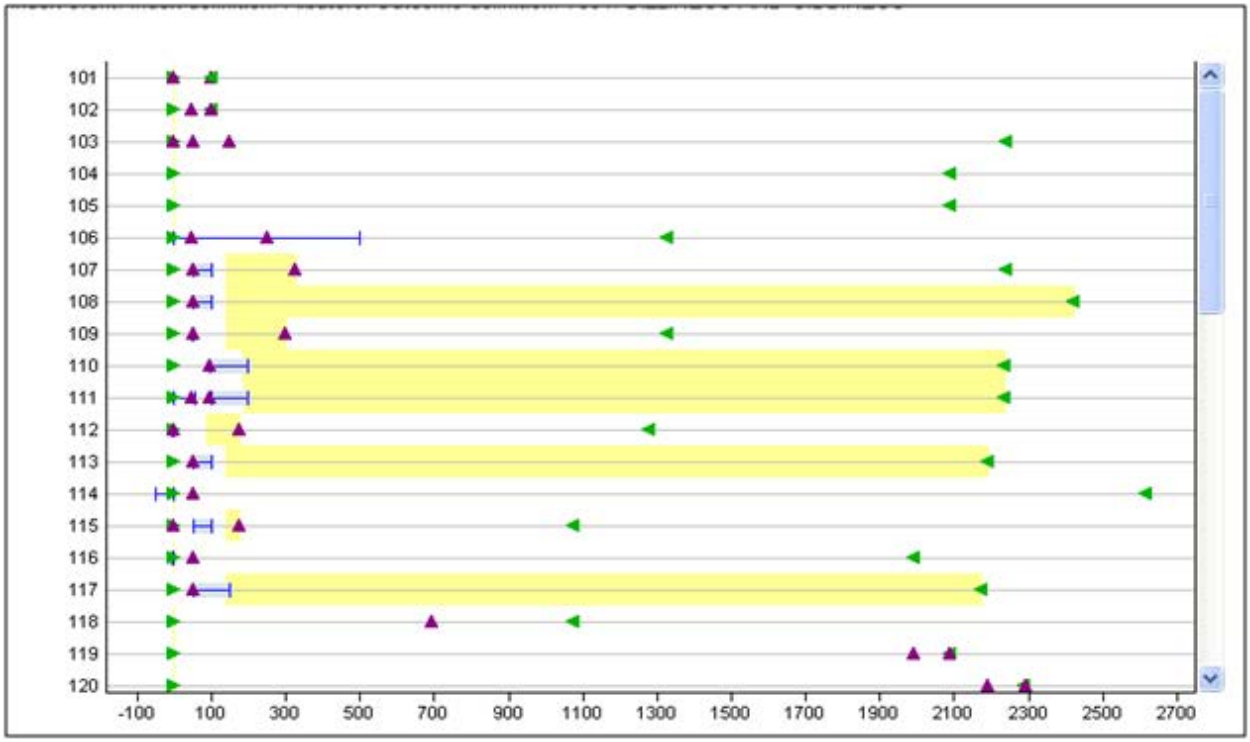
Note: When the alignment term is a drug, alignment is by prescriptions, even if the graph is configured to show only drug eras.

Option	Description
Earliest index term	(Available only for Risk-Outcome and Evaluative results.) Day 0 for a patient in the graph is the earliest occurrence of the term specified as the index event. For a drug, Day 0 is the earliest start of a prescription. If no such term exists, Day 0 is the patient's enrollment start date.
Latest index term	(Available only for Risk-Outcome and Evaluative results.) Day 0 for a patient in the graph is the latest occurrence of the term specified as the index event. For a drug, Day 0 is the latest start of a prescription. If no such term exists, Day 0 is the patient's enrollment start date.
Follow-up start	(Available only for Risk-Outcome and Evaluative results.) Day 0 for a patient in the graph is the start of follow-up as determined by the analysis.
Earliest outcome term	Day 0 for a patient in the graph is the earliest occurrence of the term specified as the outcome event. For a drug, Day 0 is

	the earliest start of a prescription. If no such term exists, Day 0 is the patient's enrollment start date.
Latest outcome term	Day 0 for a patient in the graph is the latest occurrence of the term specified as the outcome event. For a drug, Day 0 is the latest start of a prescription. If no such term exists, Day 0 is the patient's enrollment start date.
Enrollment start	Day 0 for a patient in the graph is the date of the variable designated in the data configuration as the enrollment start date.
Enrollment end	Day 0 for a patient in the graph is the date of the variable designated in the data configuration as the enrollment end date.
Analysis start	Day 0 for a patient in the graph is the first day that is within both the patient's enrollment period and the specified analysis period. Day 0 for a patient in the graph is the same as day 0 for the patient in the analysis.
Absolute time	Day 0 for a patient in the graph is 01-Jan-1970. With this option, Day 0 might not be visible on the horizontal axis.

Example

The following example shows a graph aligned by enrollment start. The enrollment starts line up at day 0 of the horizontal axis.


101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
-100
100
300
500
700
900
1100
1300
1500
1700
1900
2100
2300
2500
2700

Sorting of multi-patient timelines

In the Multi-patient Timelines Preferences dialog box, you can specify the sort order of multi-patient timelines. For more information, see Viewing and configuring multi-patient timelines.

Description

The **Sort by** option determines the order in which patients are listed on the vertical axis. Patients are listed in ascending order of the sorting option. For example, if the sorting option is Latest index term, the first patient listed is the one whose latest index term is earlier than all other patients' latest index terms.

The placement of symbols in the graph is relative to the alignment option. If you change the alignment option while keeping the sorting option the same, the sorting will likely change.

If two patients have the same relative number of days, sorting is in ascending order of patient IDs. If the alignment and sorting options are the same, the sorting is in order of patient IDs.

Sorting by index term or outcome term does not consider whether the term is within the analysis period or patient enrollment periods.

Sorting options

The following options are available for sorting.

Note: When the sorting term is a drug, sorting is by prescriptions, even if the graph is configured to show only drug eras.

Option	Description
Earliest index term	(Available only for Risk-Outcome and Evaluative results.) Sorting is by the earliest occurrence of the term specified as the index event. For a drug, sorting is by the earliest start of a prescription. If no such term exists, sorting is by the enrollment start date.
Latest index term	(Available only for Risk-Outcome and Evaluative results.) Sorting is by the latest occurrence of the term specified as the index event. For a drug, sorting is by the latest start of a prescription. If no such term exists, sorting is by the enrollment start date.
Follow-up duration	(Default selection for Risk-Outcome and Evaluative analyses.) (Available only for Risk-Outcome and Evaluative results.) Sorting is by the duration (in days) of follow-up as determined by the analysis.
Earliest outcome term	Sorting is by the earliest occurrence of the term specified as the outcome event. For a drug, sorting is by the earliest start of a

	prescription. If no such term exists, sorting is by the enrollment start date.
Latest outcome term	Sorting is by the latest occurrence of the term specified as the outcome event. For a drug, sorting is by the latest start of a prescription. If no such term exists, sorting is by the enrollment start date.
Enrollment start	(Default selection for Drug Utilization and Outcome Characterization analyses.) Sorting is by the date of the variable that is designated in the data configuration as the enrollment start date.
Enrollment end	Sorting is by the date of the variable that is designated in the data configuration as the enrollment end date.

Specifying settings

Removing the visited status for patient ID links

When you click a patient ID link, the link is marked as visited, and the color changes. The color change:

- Applies to your user name across your Empirica Healthcare Analysis sessions.
- Lasts for the number of days that you have set as your user preference in **Days to retain visited status for patient ID links** option.
- Applies across all activities within the Empirica Healthcare Analysis application.
- Applies to data associated with all configurations in the same database group.

When you remove the visited status for patient ID links, the status is removed for all the patients.

To remove the visited status for all patient ID links:

1. Click **Settings**, and click **Remove Visited Status for Patient ID Links**.

   A confirmation dialog box appears.

2. Click **OK**.

   The visited status for all patient IDs is removed. The color of the visited patient ID links is updated to a non-visited color.

Viewing the About Page

The About page contains the information about the version of the Empirica Healthcare Analysis application.

- Click **Settings**, and click **About.**

Saved lists

Saved Lists page

A saved list is a list of drugs, events, or procedures that have been grouped together and saved. You can use a saved list in an analysis run instead of specifying individual values. For example, if you have saved antipsychotic drugs in a list, you can use this list to analyze a set of antipsychotic drugs in an analysis run.

The Saved Lists page displays the following lists:

- Lists that you have created.
- Lists that you have permission to view.

For a saved list a user can assign Read/Edit/No Access permission to your entire login group or to selected individual users in your login group.

General activities

The following links and filters appear at the top of the page and affect the entire page.

- Create Saved List
- Add Lists In Bulk
- Columns
- Print
- Download
- Filter by Project or Configuration

Row-specific activities

The following menu options are available from the row menu, located in the leftmost column of the table, and affect an individual row in the table:

- View
- Edit
- Set Permissions
- Delete

Field descriptions—Saved Lists page

Field	Description
# of Terms	Number of terms in the saved list.
Configuration	Name of the configuration on which the saved list is based. For example, if the saved list is based on a run, name of the configuration used by the run.
Created By	User who created the saved list.

Created On	Date and time when the saved list was created.
Description	Description of the saved list.
ID	Identifier that was automatically assigned to the saved list when it was created. The saved list IDs are unique and are not reused if they are deleted.
List Type	Type of variable to which the saved list is attached.
Modified	Date and time when the saved list was last modified.
Modified By	User who modified the saved list.
Name	Name of the saved list.
Project	Name of the project associated with the saved list.
Variable	Name of the variable for which the saved list contains terms.

Creating a saved list

1. Click **Settings**, and click **Manage Saved Lists.**

   The Saved Lists page appears.

2. Click **Create Saved List**.

   The Saved List dialog box appears.

3. Fill in the fields:

   A. Type a **Name** and **Description**.

   B. Select a **Project** and **Configuration.**

   C. In the **Associated Variable** drop-down list, select a variable.

   D. Add values to the list. For more information, see Selecting values.

   E. To verify that the specified values exist in the source data, select **Validate list when saving**. If any of the values does not have a matching value in the source data, an error message appears when you save the list, and the unmatched values appear at the end of the list.

4. Click **Save**.

   The list is created and appears on the Saved Lists page. The permissions assigned to the list determine whether other users can use it.

Setting permissions for a saved list

You can set permissions for a saved list that you have created. Additionally, if you have the Edit permission, you can use, edit, modify permissions for, and delete the saved list.

If you have the Read permission for a saved list, you can use the saved list.

1. Click **Settings**, and click **Manage Saved Lists**.

   The Saved Lists page appears.

2. Select the row menu (), and select **Set Permissions**.

   The Saved List Permissions dialog box appears.

3. Grant permissions:

   - To grant permissions to other users for your login group, select **Read** or **Edit** next to the Group Permissions. By default, a group has **No Access** permission.
   - To grant permissions to the individual users in your login group, scroll to the Individual Permissions area, and click **Read** or **Edit** for a user.

   Note: User permissions for a saved list are determined by the settings for your user name and login group. For example, if you have Read access but your login group has Edit access, you also have Edit access.

4. After you finish setting permissions, click **Save**.

Users

User administration terms

Term	Definition
User	A person with a user name and password for logging in to the application.
User profile	Attributes, such as a user role, permissions, and preferences, which you can apply to one or more users. User profiles are similar to templates and allow you to set up users quickly. If you set user preferences by applying a user profile to users, each user can override the preferences.
User role	A set of permissions. You can assign one or more roles to each user or user profile.
Permission	A right that allows a user to perform an activity. You can assign permissions to users, user profiles, and roles.
Login group	A group of users.

Viewing a list of users

- Click **Settings**, and click **Edit Users**.

  The Users page appears. The following information is provided about each user.

Column	Description
Authentication	**Local**, indicating the standard authentication method for the application.
Email	Email address associated with the user. The address (or addresses, separated by commas) is the default notification address for analysis runs if the run creator chooses email notification. This address is also used when a user with appropriate permissions sends a message to all users.
Name	Full name of the user.
Quota	Maximum amount of server space in megabytes (MB) that the user is permitted to use for creating runs. **Unlimited** appears for users without a quota.
SSO Login	**Enabled** appears if single sign-on authentication is enabled for the user. **Disabled** appears if local authentication is used.
Status	**Enabled** or **Disabled**, indicating whether the user can log in.
User Name	Name of the user.

Related Topics

Adding and editing a user

Deleting a user

Renaming a user

Changing a password

Changing a password

You can edit only users in your login group.

Prerequisite

- You must have the **Administer Users** permission to change the password of another user in your login group.

To change a password:

1. Click **Settings**, and click **Edit Users**.

   The Users page appears.

2. Select the row menu (), and select **Edit**.

   The Edit User page appears.

3. Click **Change Password**, located at the bottom of the page.

   The Change Password page appears.

4. Fill in the fields. The site administrator sets the password requirements, including length and required characters and whether you can reuse old passwords.

5. Click **Change Password**.

   The password is updated. The password expiration period, if one has been set, begins for the new password. The next time the user logs in, the user must use the new password.

Adding and editing a user

1. Click **Settings**, and click **Edit Users**.

   The Users page appears.

2. To add a user, click **Add a New User**.

   If you are going to enable Single Sign-on login for a user, you must also create and configure the user in the Oracle Identity Manager.

   To edit a user, select the row menu (), and select **Edit**.

3. Fill in the fields according to the table below, and click **Save**.

Field descriptions—Add User and Edit User page

Field	Description	Note
Authentication	**Local**, indicating that standard authentication applies.	Read-only.
Username	Unique name of the user account. The application converts usernames to lowercase and thus does not allow duplicate usernames that differ only in case. For example, if you create user **jkelley**, you cannot create a user **JKelley**.	Appears only when you create a user. Maximum length: 100 characters. You can reuse deleted usernames.
First Name Last Name	Name of the user.	Maximum length: 64 characters for each field.
Email	Email address of the user.	Separate multiple addresses using a comma. Oracle recommends adding an email address for each user.
User Profile	User profile to apply to the user.	
Quota	Maximum amount of server space in megabytes (M) that the user is permitted to use for creating runs. If this limit is exceeded, the user cannot submit new runs or reruns. To give a user unlimited storage space, leave the field blank. To prevent the user from creating any analysis runs, even if the user has the	

	appropriate permissions, type 0.	
Login Group	Name of the login group to which the user belongs.	
Password Confirm Password	Password for logging in. You must create passwords according to the password restrictions set by your site administrator. Follow recommendations by your organization related to creating secure passwords.	Maximum length: 64 characters for each field. Users can modify their passwords.
Password never expires	When selected, the user's password never expires. Oracle recommends creating at least one user with a password that does not expire.	
User must change password at next login	When selected, the user must change the password at the next login.	This option is deselected after the user changes the password.
Account locked	A user cannot log in until an administrator deselects the option.	This option is selected when a user's password expires, or when a user types an incorrect login password the maximum number of times. When a user's password expires, the user is prompted to change the password.
Account disabled	A user cannot log in until an administrator deselects the option.	Only an administrator can select and deselect this option.
Enable SSO Login when SSO is configured	When selected, single sign-on is enabled for the user.	You must have completed all single sign-on configuration tasks before selecting this check box.

Related topics

Assigning a role to a user

Renaming a user

The application converts user names to lowercase automatically upon creation, and thus does not allow duplicate user names that differ only in case. For example, if you create user jkelley and then attempt to create user JKelley, an error occurs.

After you change a user name:

- The original user name can no longer be used to log in.
- The new user name appears throughout the application, including historical activities in the User Activity Audit Trail.
- You can create another user with the original user name.

To rename a user:

1. Click **Settings**, and click **Edit Users**.

   The Users page appears.

2. Select the row menu (), and select **Edit**.

   The Edit User page appears.

3. At the bottom of the page, click **Rename User**.

   The Rename User page appears.

4. Type a new user name. The name must be unique with a maximum of 100 characters.

5. Click **Save**.

   If the affected user is currently logged in, the user can continue working; however, after the user logs out, the user must log in using the new user name.

Related topics

Adding and editing a user

Changing a password

Deleting a user

After you delete a user:

- The user can no longer log in; however, the user name is not removed from the database.
- The only place where the user name is visible is the User Activity Audit Trail, which allows you to view activities performed by the deleted user. The user name is followed by **(deleted)**.
- You can create another user with the same user name as the deleted user.

To delete a user:

1. Click **Settings**, and click **Edit Users**.

   The Users page appears.

2. Select the row menu (), and select **Delete**.

   You are asked to confirm the deletion.

3. Click **Yes**.

   The user is deleted. If the deleted user is logged in, that user can continue working; however, after the user logs out, the user cannot log in again.

Permissions and roles

User permissions

Permissions can be assigned to users or user roles. Several predefined user roles are available, and you can create additional roles as needed.

Note: A user's ability to perform activities on an object might depend on whether the object is published and whether the user has the Administer Users permission. For example, users can view the results of analysis runs that they created or that have been published to their login group. If users have the Administer Users permission, they can view published and unpublished runs created by any user in their login group.

Safety Data permissions

Permission	Activities
View Raw Risk Database	• Drill down to view patient details from the drilldown menu that appears when you click a patient count in analysis results. • View single-patient and multi-patient timelines. **Note**: Oracle recommends giving this permission to most users so that they can drill down to view patient details.
Download Raw Data	Download patient details from the drilldown menu, which appears when you click a patient count in analysis results and click a graph of analysis results.

Analysis Runs permissions

Permission	Activities
View Analysis Results	• View results of an analysis run. • Create, edit, or delete a saved list.
Create Analysis Run	• Create an analysis run. • Re-run an analysis run. • Rename, cancel, or delete an analysis run that you created. • View jobs for a run.
Save Intermediate Files	Save intermediate data files created during an analysis run.
Access Advanced Run Options	Specify advanced parameters for an analysis run.

Delete Any Run	Cancel or delete analysis runs (published or unpublished) created by other users in your login group.

Queries and Reports permissions

Permission	Activities
Browse Data	Use the Browse Data feature.
Create Queries/Patient Groups	If the user also has the View Raw Safety Database permission: • Create, copy, edit, rename, or delete a query or patient group. • Transfer patients to an existing patient group.
Create Report Definitions	Create, copy, edit, or delete report definitions. No permission is required to run report definitions.
Create Report Outputs	Create, edit, copy, or delete report outputs No permission is required to view report outputs.

Administration permissions

Permission	Activities
Administer Users	• View, add, edit, or delete user names and user profiles in your login group. • Send a message to users in the same login group. • Arrange table columns and set that layout as the default for your login group. • View the User Activity Audit Trail for users in the same login group. • View currently logged in users in the same login group. **Notes:** • Users with this permission can edit their own login group but cannot create or delete login groups. • For some activities, users with this permission can act on objects (including unpublished objects) created by users in the same login group.
Manage Configurations	• View data configurations. • Create, import, copy, validate, or edit a data configuration for which you have the Edit permission. • Delete a data configuration that does not have existing runs, patient groups, saved lists, report definitions, or report outputs associated with it.

- View the Analysis Setup tab and create patient statistics runs.

User roles

The following roles are available for all installations.

Role	Permissions
Analysis Run Reviewer	View Raw Safety Database Download Raw Data View Analysis Results Browse Data Create Queries/Patient Groups
Analysis Run Designer	View Raw Safety Database Download Raw Data View Analysis Results Create Analysis Run Save Intermediate Files Browse Data Create Queries/Patient Groups
Customer Administrator	Administer Users
Report Designer	View Raw Safety Database Browse Data Create Queries/Patient Groups Create Report Definitions Create Report Outputs
Report Producer	View Raw Safety Database Browse Data Create Queries/Patient Groups Create Report Outputs

Creating and editing a role

For a list of predefined roles, including their permission, see User roles.

Prerequisites

- You must be a superuser.

To create a role:

1. Click **Settings**, and click **Edit Roles**.

   The Edit Roles page appears, displaying all roles.
2. Click **Create New Role**.

   The Create New Role page appears.
3. Type the name of the role, and click **Save**.

   The Edit Roles page appears.
4. Select the permissions to assign to the role.
5. Click **Save**.

To edit a role:

1. Click **Settings**, and click **Edit Roles**.

   The Edit Roles page appears, displaying all roles.
2. To edit a role, locate the role in the list, and click **Edit** next to it.

   The Edit Roles page appears.
3. Select the permissions to assign to the role.
4. Click **Save**.

Assigning a role to a user or user profile

The application is delivered with predefined user roles.

When you create or edit a user and user profile, you can assign one or more user roles to the user or user profile. You can also assign individual permissions to the user or user profile.

1. To assign a role to a user, click **Settings**, and click **Edit Users**.

   To assign a role to a user profile, click **Settings**, and click **Edit User Profiles**.

   Either the User or User Profiles page appears.

2. Select the row menu (), and select **Edit**.

3. At the bottom of the page, click **Assign Roles**.

   The Assign Roles page appears, listing all user roles.

4. Select or deselect user roles to assign roles to or remove roles from the user or user profile.

5. Click **Save**.

   The user or user profile is given the assigned roles and their associated permissions.

   If you modified a user, and the user is logged in, the user is not affected by the changes until the next time the user logs in.

Assigning permissions to a user or user profile

If a user or user profile has permissions from the assigned role, that information is indicated on this page. You can assign additional permissions to the individual user or user profile; however, you cannot disable permissions that have been assigned via a user role.

1. To assign permissions to a user, click **Settings**, and click **Edit Users**.

   To assign permissions to a user profile, click **Settings**, and click **Edit User Profiles**.

   Either the User or User Profiles page appears.

2. Select the row menu (), and select **Edit**.

3. At the bottom of the page, click **Assign Permissions**.

   The Assign Permissions page appears, listing all permissions.

4. Select or deselect permissions to assign permissions to or remove permissions from the user.

5. Click **Save**.

   The user or user profile is given the permissions.

   If you modified a user and the user is logged in, the user is not affected by the changes until the next time the user logs in.

User profiles

Adding and editing a user profile

1. Click **Settings**, and click **Edit User Profiles**.

   The User Profiles page appears.

2. To add a user profile, click **Add a New User Profile**.

   To edit a user profile, select the row menu (), and select **Edit**.

3. Fill in the fields according to the table below, and click **Save**.

Field descriptions—Add User Profile and Edit User Profile page

Field	Description
User Profile Name	(Appears only when you are adding a user profile.) Name of the user profile.
Quota	Maximum amount of server space in megabytes (M) that a user created from the user profile is permitted to use for creating runs. If this limit is exceeded, the user cannot submit new runs or reruns. To give a user unlimited storage space, leave the field blank. To prevent the user from creating any analysis runs, even if the user has the appropriate permissions, type 0.
Enable SSO login when SSO is configured	When selected, single sign-on is enabled for a user created from the user profile. You must have completed all single sign-on configuration tasks before selecting this check box.

Login groups

Creating and editing a login group

1. Click **Settings**, and click **Edit Login Groups**.

   The Login Group Settings page appears.

2. To create a login group:

   A. Click **Create New Login Group**.

      The Create Login Group page appears.

   B. Type a name for the login group, and click **Save**.

      The Login Group Settings page appears.

   C. Fill in the fields according to the following dialog box, and click **Save**.

3. To edit a login group:

   A. Next to the login group to edit, click **Edit**.

      The Login Group Settings page appears.

   B. Fill in the fields according to the following dialog box, and click **Save**.

Field descriptions—Login Group Settings page

Field	Description	
Logo Image	Image file that appears as the product logo. The software scales a logo image that is not exactly 50 pixels in height to fit within the space allotted for the header. The aspect ratio is maintained so that the width of the image is also scaled.	The file must be an image file and must be in the \image subdirectory of the server location to which the Empirica Healthcare Analysis software was installed.
Home Page	Name of the file that contains the customized HTML for the Home page.	The file must be an HTML file in the \customhomes subdirectory of the server location to which the Empirica Healthcare Analysis software was installed.

Specifying the Home page for a login group

You specify a Home page for each login group. The default Home page is named home.inc.

Prerequisites

- An administrator must create the Home page HTML file.
- The HTML file must exist in the \customhomes subdirectory of the server location in which the Empirica Healthcare Analysis software was installed.

To specify the Home page for a login group:

1. Select **Settings > Edit Login Groups**.

   The Login Group Settings page appears.

2. In the Home Page field, type the name of the file that contains the customized HTML for the Home page.

3. Click **Save**.

   Users who are associated with the edited login group see the changes the next time they log in.

Administering the software

Setting site options

A site option is a setting that customizes an aspect of the software for all users.

Prerequisites

- You must be a superuser.

To set site options:

1. Click **Settings**, and click **Set Site Options**.

   The Site Options page appears.

2. Fill in the fields according to the following table, and click **Save**.

Field descriptions—Site Options page

Field	Description
Expiration	Number of days after which passwords expire. When a password expires, a user has the opportunity to create a new password. This optional setting affects passwords for existing and new users. The expiration period is counted from when the user was created or when the password was last changed. **Note:** When you add or edit a user, you can specify that the user's password never expires.
Expiration Warning	Number of days prior to password expiration that a warning message appears when a user logs in. This optional setting affects passwords for existing and new users. The expiration warning period is counted from when the user was created or when the password was last changed. For example, if you type 5, then five days before a password is due to expire, a warning message appears when the user logs in. The user can either change the password or continue without changing the password. If the user does not change the password, the message continues to appear upon login, counting down the number of days until password expiration.
Minimum Length	Minimum length of new passwords. Passwords can be up to 64 characters long. Existing passwords are not affected by this value.

Number of Attempts Allowed	Number of times that a user can attempt to log in with an incorrect or missing password before the account is disabled. When the limit is reached, an email message is sent to all superusers and to all user names with the Administer Users permission in the same login group as the disabled user name. **Note:** When a user account is disabled, the **Account disabled** checkbox is selected for the user. To enable the user account, edit the user and deselect the checkbox.
Number of Passwords Retained	Number of unique passwords that are retained in the user's history. Passwords retained in a user's history cannot be reused. Each time a user changes a password, the old password is retained, up to the specified number of passwords, such as 10. A new password cannot match any passwords in the history. To allow password reuse, set this value to 0. To prevent password reuse, set this value to a high number, such as 1000.
Minimum number of each of the following characters required in new passwords	Requirements on password content. The number in each field determines the number of each type of character that a password must contain.
SMTP Server	Name of the SMTP server to use for email notification of run completion, feedback, and errors. This field must have an accurate server name for the email functions to work.
Feedback Email	Email address to which users send feedback. Separate multiple addresses with commas.
Error Email	Address to which an email message is sent when a system error occurs. Separate multiple addresses with commas. **Note:** System errors also appear in the user activity audit trail, which frequently includes details about the error.
Date Format Time Format	Date and time format to use for server datetimes that are displayed in the software. **Note:** UTC is Coordinated Universal Time.

Log Level	Type of information to include in the weberror.log file on the server: • Error • Warning (default) • Information • Debug
Max Memory Per Report	Maximum amount, in MB, of memory that a single report can use when executing. Typically, you choose a value that is no larger than 768 MB. Before changing this option, consult Oracle. **Tip:** To decrease the amount of memory used by reports, edit report attributes and set the **Generate Drilldown Information** option to **No**.
Number of patients per MP timelines page	Maximum number of patients to appear by default on each page of a multi-patient timelines graph. The value must be from 1 to 500. Users can override the value.
Max patients in SP timelines	Maximum number of patients that can be included in single-patient timelines.
Max patients in MP timelines	Maximum number of patients that can be included in multi-patient timelines.
Allow Patient Comment/Review/Exclusion	When selected, the Reviewer Input section appears on the Patient Details page. This option affects only the Patient Details page that you open from the Patient Queries tab.
Show "Patient Queries" tab	When selected, the Patient Queries tab appears.
Show "Descriptive Analysis" tab	When selected, the Descriptive Analysis tab appears if users have the appropriate permissions.
Show "Evaluative Analysis" tab	When selected, the Evaluative Analysis tab appears if users have the appropriate permissions.
Show "Reports" tab	When selected, the Reports tab appears.
Enable Query Wizard	When selected, the Query Wizard is available for defining queries.
Enable Browse Data	When selected, the Browse Data feature is available if users have the appropriate permissions.
Enable Second Level Drilldown from Patient	When selected, users can drill down to patient details. When deselected, no users can drill down to patient details,

Counts	regardless of their permissions.
Enable Download Patient Details	When selected, users can download patient details if they have the appropriate permissions and the configuration option for downloading patient details.
Enable User Preference to display Patient Timelines as applet	The **Display Patient Timelines as applet** preference is available.
Add message to Notes for Results (Table and Graphs)	Text to appear below tables of analysis results. For example, the message might read: These data do not, by themselves, demonstrate causal associations; they may serve as an impetus for further investigation.

Sending a message to all users

If you have the Administer Users permission, you can send an email message to all users who meet the following criteria:

- Have an email address associated with their user name.
- Are in your login group.
- Have the **Account disabled** setting on the Add/Edit User page deselected, indicating an active user account.

Note: Do not include private data or PII data in a message.

To send a message to all users:

1. Click **Settings**, and click **Send Message to All Users**.

   The Send to All Users page appears.

2. Fill in the fields, and click **Send**.

Restarting the listener process

The listener process identifies background work and assigns it to processes for execution. The listener process starts when the software starts. The process updates an associated heartbeat date and time value every 2 to 30 seconds while the software is running.

Occasionally, the listener process terminates unexpectedly, either because the process encounters an error or is stopped by a system operator.

When a run is submitted, the software checks the heartbeat value to determine if the listener process is still running. If the listener process is not running, the software restarts the process automatically.

Note: You can restart the listener process manually at any time.

Prerequisites

- You must be a superuser.

To restart the listener process:

- Click **Settings**, and click **Restart Listener**.

Monitoring the software

Viewing a list of users who are logged in

You can see only users in your login group.

- Click **Settings**, and click **View Currently Logged In Users**.
  The Current Users page appears and displays all users who are logged in.

Viewing an audit trail of user activity

You can see the activity of only users in your login group.

1. Click **Settings**, and click **View User Activity Audit Trail**.

   The User Activity Audit Trail page appears.

2. Select a user and optionally, start and end dates.

3. Select checkboxes for the types of user activities that you want to see.

4. Click **View User Activity**.

   An audit trail appears for the selected activities performed by the user.

Viewing the status of the server

The following types of computer processing are supported:

- Foreground or interactive processing, such as the specification of runs.
- Background or batch processing, such as the execution of runs.

If you are using a single-processor server, the processor must perform both kinds of processing. If you are running on a multi-processor server, you can control the way the available processors are used to perform foreground or background processing.

On the Server Status page, you can set the maximum number of processors to use for background processing; that is, the number of concurrent runs that can be performed. For a dedicated server, Oracle recommends that you set the maximum number of processors to one less than the number of processors that are on the server. For example, if a dedicated server has three processors, and you set the maximum number of processors to two, two runs can run at the same time.

Prerequisites

- You must be a superuser.

To specify the maximum number of processors:

1. Click **Settings**, and click **View Server Status**.

   The Server Status page appears.

2. Change the value in the **Maximum Processors** field, and click **Save**.

   A message tells you that your update was successful.

3. Click **Continue**.

Viewing the free space in a tablespace

The option to view free disk space is for users who are familiar with Oracle storage mechanisms. This feature is useful for monitoring the Max Utilization level of a tablespace. When the level approaches a threshold predefined by your installation, such as 75 percent, you can allocate additional data files.

This option displays information for tablespaces used to store system data created and managed by the Empirica Healthcare Analysis application. System data includes user-related information, run definition properties, run outputs, patient lists, reports, and so on.

To review the utilization of tablespaces used to store source data, you must use Oracle DBA tools.

Prerequisites

- You must be a superuser.

To view disk space:

1. Click **Settings**, and click **View Free Space**.

   The View Free Space page appears.

2. From the drop-down list, select an Oracle tablespace.

   Information about remaining space for each file in the selected tablespace appears.

Configurations and patient statistics runs

About configurations

Configurations

A configuration makes source data available to end users and includes the following information:

- Attributes, such as the names of the configuration and the Oracle account where the configuration resides.
- A list of variables that have their own attributes, such as type and subtype.

Variables reference tables in the source data account. For example, demographic variables map to source demographic tables; drug variables map to source drug tables; and event (diagnosis) variables map to source event tables. If source data includes procedure data, procedure variables can map to source procedure variables.

Configurations determine the following:

- The drug variables, event variables, and procedure variables available for analysis or querying.
- The columns in the source data used to determine patient enrollment start and stop dates.
- The variables available for defining breakdowns or covariates in analysis runs.
- The hierarchies of terms associated with drugs or events.
- That patient information available in drilldown displays.
- The variables available in the Browse Data feature.
- The variables available for HDPS dimensions, if they are different from the default variables.

A configuration is available to end users after an administrator validates it and grants permissions to it.

Multiple configurations

You can set up multiple configurations that point to the same source data, with each configuration representing source data in a different way.

For example, suppose that the source data includes the following information:

- The date on which a prescription was written.
- The date on which the prescription was filled.

You can set up a configuration with **date written** as the start of prescriptions and another configuration with **date filled** as the start of prescriptions.

You must inform end users of the configurations they should use for various purposes and give the users access to appropriate configurations.

Ancillary tables

In addition to referencing source data tables, a configuration or its variables might reference ancillary tables. For example, a drilldown map table that is referenced by the configuration determines the data that appears when end users drill down on patient counts.

Hierarchies of terms

The Empirica Healthcare Analysis software supports the use of a dictionary, thesaurus, or other standardized terminology that organizes or classifies terms for drugs or events. The software stores the terms and their relationships in a special Oracle account. You associate the terms with drug or event variables in a configuration.

If hierarchy information is set up, end users can select drug or event terms from the hierarchy as an alternative to using the Select Available Values link. To support a hierarchy of terms, complete the following steps:

1. Set up the special Oracle account containing the terms, if it does not exist.
2. Specify the drug hierarchy account and event hierarchy account for a configuration.
3. In the configuration, include a variable for each level of the hierarchy. Specify the selection type for the variable.

Configuration tables

Each data account contains a table that stores attributes of configurations for the account.

For each configuration in the account, there is a table that stores attributes of variables for the configuration.

Validation checks for configurations

The validation process for configurations checks the configuration attributes and variable attributes and verifies whether the following statements are true:

- The configuration and variables are set up correctly.
- The source data does not contains certain invalid characters.

Validation checks can result in warning or error messages. If error messages occur, you must correct the error messages before you can validate the configuration.

Imported configurations

If you use a script to create a configuration and import the configuration without visiting the pages for specifying attributes of a configuration or variables, the validation process performs the checks that would normally occur on the pages.

Validation checks that produce errors

General validation checks

An error occurs for each of the following requirements that are not met:

- The configuration must have variables defined for it.
- Each source data table referenced by the configuration must have one and only one variable with the **Patient ID** type.
- The drilldown map table must have a row for the LIST section.
- Source data must not contain any of the following characters:
  - Backslash (\).
  - Back quote (').
  - Double quote (").
  - Exclamation point (!).
  - Dollar sign ($).
  - Vertical bar (|).
  - At sign (@).
  - Certain non-printable characters, such as carriage returns.

  If the source data contains any of the characters, an error message occurs for each variable that has invalid values. For information about the specific invalid values, view the INVALID_SOURCE_DATA table in the data account. You must correct the errors in the source data before the configuration can be validated successfully.

Validation checks for the demographics table

An error occurs for each of the following requirements that are not met:

- For the demographics table named in the source description table, at least one variable must be defined.
- The configuration must have one and only one variable of each of the following types and subtypes:

- **Other** type and the **Enrollment Start Date** subtype.
- **Other** type and the **Enrollment End Date** subtype.

Validation checks for the drug table

An error occurs for each of the following requirements that are not met:

- The source description table specified for the configuration must have a drug table defined.
- For the drug table named in the source description table, the configuration must have at least one variable with the **Drug** type.
- For the drug table named in the source description table, there must be one and only one variable with each of the following types:
  - **Record Date**
  - **Record Duration**
- There cannot be more than one variable with the **Other** type and **Drug Label** subtype.

Validation checks for the event table

An error occurs for each of the following requirements that are not met:

- The source description table specified for the configuration must have an event table defined.
- For the event table named in the source description table, the configuration must have at least one variable with the **Event** type.
- For the event table named in the source description table, there must be one and only one variable with the **Record Date** type.
- There cannot be more than one variable with the **Other** type and **Event Label** subtype.

Validation checks for the procedure table

An error occurs for each of the following requirements that are not met:

- If the source description table references a procedure table, the procedure table must contain at least one variable with the **Event** type and the **Procedure** subtype.
- If the procedure table contains one or more variables with the **Event** type and the **Procedure** subtype, the source description table must reference a procedure table.
- For the procedure table named in the source description table, there must be one and only one variable with the **Record Date** type.
- There cannot be more than one variable with the **Other** type and **Proc Label** subtype.

Validation checks that produce warnings

A warning occurs for each of the following requirements that are not met:

- The drilldown map table specified for the configuration should include one row for each of the following sections:
  - CASE

  - DRUG
  - EVENT
- The configuration should have a variable with the **Gender** type.
- The configuration should have a variable with the **Age** type and either the **Age at enroll-start** or **Age at enroll-end** subtype.
- The configuration should have one and only one variable with the **Other** type and **Drug Label** subtype.
- The configuration should have one and only one variable with the **Other** type and **Event Label** subtype.
- Source data should not have leading or trailing white spaces.

Additional activities

In addition, the validation process performs the following steps:

- Create a unique values table for any variable that does not have an associated unique values table and that has the selection type **Distinct Value List**, **Select From Drug Hierarchy**, or **Select From Event Hierarchy**.
- Create a new patient ID table for any variable for which one is requested.
- Create Oracle column indexes for columns referenced by any variable, if the indexes do not already exist.

  If a variable references a column in a view, an index is not added.

Manage Configurations page

On the Manage Configurations page, you view existing configurations, validate configurations, and add new configurations.

General activities

The following links and filters appear at the top of the page and affect the entire page.

- Add Configuration
- Import Configurations
- Validate Configurations
- Columns
- Print
- Download
- Select Rows

Row-specific activities

The following menu options are available from the row menu, located in the leftmost column of the table, and affect an individual row in the table. Your permissions determine the options that appear.

- Edit

Field descriptions—Manage Configurations page

Field	Description
Account	Name of the Oracle account in which the configuration resides.
Database Group	Name of the database group to which the configuration is assigned. The application uses the database group when determining color changes for patient ID links that are visited by end users.
Desc	Description of the configuration.
Drilldown Map	Name of the drilldown map table, which specifies the source of patient details that appear when an end user drills down on patient counts or patient IDs.
Is Valid	Yes or No, indicating whether the configuration is valid. Newly created, copied, or imported configurations are invalid. After a configuration is validated, its status either becomes valid or remains invalid. A valid configuration becomes invalid if any of the following occurs: • The configuration fails validation. • You click **Save** on the Edit Configuration Details page, where you

edit a configuration.

- You click **Save** on the Edit Variable page, where you edit a variable.

Name	Name of the configuration.
Owner	Name of the user who created the configuration.
Source Desc	Name of the source description table, which identifies the source database tables.
Table	Name of the Oracle table containing attributes of the variables.

Modify Configurations page

On the Modify Configurations page, you can modify and validate a single configuration.

General activities

The following links and buttons appear at the top of the page and affect the configuration you are editing.

- Validate Now
- Validate Later

> **Note:** Both Validate buttons appear regardless of the validation status of the configuration.

- Add New Variable
- Copy this Configuration
- Delete this Configuration
- Edit (located in the last column of the first table)
- Columns
- Print
- Download

Row-specific activities

The following menu options are available from the row menu, located in the leftmost column of the table, and affect a variable in the configuration you are editing.

- Edit
- Delete

Field descriptions—Modify Configuration page, top table

The following fields provide information about the configuration.

Field	Description
ID	Unique identifier that was assigned to the configuration.
Name	Name of the configuration.
Description	Description of the configuration.
Database Group	Name of the database group to which the configuration is assigned. The Empirica Healthcare Analysis software uses the database group when determining color changes for patient ID links that are visited by end users.
Owner	First and last name of the user who created the configuration.

Table	Name of the configuration table, which includes variables and their attributes and is in the data account.
Drilldown Map Table	Name of the drilldown map table, which specifies the source of patient details that appear when an end user drills down on patient counts or patient IDs.
Source Description Table	Name of the source description table, which identifies the source database tables.
Drug Hierarchy Account	Name of the Oracle account containing the hierarchy for variables with the selection type **Select From Drug Hierarchy**.
Event Hierarchy Account	Name of the Oracle account containing the hierarchy for variables with the selection type **Select From Event Hierarchy**.
Valid	Yes or No, indicating whether the configuration is valid. Newly created, copied, or imported configurations are invalid. After a configuration is validated, its status either becomes valid or remains invalid. A valid configuration becomes invalid if any of the following occurs: • The configuration fails validation. • You click **Save** on the Edit Configuration Details page, where you edit a configuration. • You click **Save** on the Edit Variable page, where you edit a variable. • You import configurations for the data account in which the configuration resides.

Field descriptions—Modify Configuration page, bottom table

The following fields provide information about each variable in the configuration.

Field	Description
Column	Name of the database table column that corresponds to the variable.
Desc	Description of the variable.
Hierarchy Level	Hierarchy level, if any, specified for variables with a selection type of **Select From Drug Hierarchy** or **Select From Event Hierarchy**.
Patient ID Table	Name of the patient ID table for the variable.
Prefix	Prefix of the variable.
Selection Type	See Variable selection types.

Table	Name of the source database table or view containing the column that corresponds to the variable.
Unique Values Table	Name of the unique values table, if any, associated with the variable. For more information, see Selecting a unique values table for a variable.
Variable	Name of the variable.
Variable Type	See Variable types. If the variable has a subtype, the subtype appears in parentheses. For more information, see Variable subtypes.
Visibility	• **Analyses and Queries** appears if the variable is available for analysis runs and queries. For the variable, Yes is selected for Hide From Query Wizard. • **Analyses** appears if the variable is available for analysis runs but not for queries. For the variable, No is selected for Hide From Query Wizard. **Note:** You build queries using the query wizard on the Patient Queries page.

Working with configurations

Adding a configuration

Prerequisites

- You must have the Manage Configurations permission or be a superuser.

To add a configuration:

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Click **Add Configuration**.

   The Add/Import Configurations page appears, listing the names of accounts that contain added or imported configurations.

3. Select the account where data resides for the configuration that you want to add:

   - If the account appears in the list, select it.
   - If the account does not appear, click **Choose a Different Account**.

   You are prompted to type a name.

4. Type the name of the Oracle account for the configuration, and click **OK**.

   The Add Configuration Page appears.

5. Select a drilldown map table, which is used for drilldown on patient counts and patient IDs, for the configuration:

   A. Next to the **Drilldown Map** field, click **Select/Edit Table**.

      The Select/Edit Drilldown Table dialog box appears.

   B. From the **Select Drilldown Table** drop-down list, select a table. The list includes all tables that are in the data account and that have the appropriate structure.

      Values for the selected drilldown table appear in the dialog box.

   C. Click **OK**.

   Note: To create or edit a table, rather than select an existing table, see Creating, copying, and editing a drilldown map table.

6. Select a source description table, which is used to identify the source database tables, for the configuration:

   A. Next to the **Source Description Table** field, click **Select/Edit Table**.

      The Select/Edit Source Description Table dialog box appears.

B. From the **Select Source Description Table** drop-down list, select a source description table. The list includes all tables that are in the data account and that have the appropriate structure.

   Values for the selected table appear in the dialog box.

C. Click **Save**.

Note: To create or edit a table, rather than select an existing table, see Creating and editing a source description table.

7. To select an HDPS dimensions table, which displays custom dimensions in certain evaluative analysis runs:

   A. Next to the **HDPS Dimensions Table** field, click **Select Table**.

      The Select Table dialog box appears.

   B. Select a table that has been set up correctly.

   C. Click **Save**.

8. Fill in the remaining fields according to the following table, and click **Save**.

   The configuration is created with a status of invalid.

9. Assign permissions to the configuration.

Field descriptions—Add Configuration Page

Field	Description
Name	Unique name of the configuration, up to 100 characters. The name can include the following characters: • Letters • Digits • Blank space • Plus sign (+) • Hyphen (-) • Underscore (_) • Parentheses () • Multi-byte characters If a configuration was deleted by reference only, the configuration is still in the database, and you cannot add a configuration with the same name.
Description	Description that appears to end users when they browse for configurations.
Drug Hierarchy	Name of the Oracle account containing the hierarchy for variables with

Account	the selection type **Select From Drug Hierarchy**.
Event Hierarchy Account	Name of the Oracle account containing the event hierarchy of terms for variables with the **Select From Event Hierarchy** selection type.
Configuration Table Name	Name of the table, which is created when you add the configuration, that will contain the variable attributes in the Oracle account to which you are connected. The name must be unique in the Oracle account and must be a valid Oracle table name.

Assigning permissions to a configuration

- If you have the Manage Configurations permission, you can set permissions for your login group and users in your login group.
- If you are a superuser, you can set permission for any user or login group.

To assign permissions to a configuration:

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Select the row menu for a configuration (), and select **Edit**.

   The Modify Configuration page appears.

3. In the last column of the first table, click the **Edit** link.

4. Set permissions for login groups and users:

   A. Scroll to the Group Permissions and Individual Permissions tables.

   B. Set permissions for each login group and individual user by selecting one of the following options. A user's permission for a configuration is the broadest permission for the combination of the user name and login group.

      - **No Access**
      - **Read**—The configuration can be selected and used for activities such as analysis.
      - **Edit**—The configuration can be selected and used for activities such as analysis. Users with the Manage Configurations permission can also work with the configuration on the Manage Configurations page.

   Note: The configuration permissions (No Access, Read, Edit) are ignored for a superuser, who can always select a configuration for end user activities and can always see all configurations on the Manage Configurations page.

   C. To allow a user or login group to download patient details, select **Allow Details Download** for the user or login group. The user or login group must also have the appropriate permissions, and the site option must also be enabled.

   If the configuration was valid, it becomes invalid, and you must validate it again.

Importing configurations

When you import configurations, you import all configurations from a data account. A configuration is in a data account if it meets any of the following guidelines:

- The configuration was created outside the application.

  For example, data preparers might have used a script to create a configuration table.
- The configuration was deleted from the application by reference only.

To import configurations:

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Click **Import Configurations**.

   The Add/Import Configurations page appears, listing the names of accounts containing added or imported configurations.

3. To import configurations from a listed account, click the account.

   To import configurations from a different account, click **Import Configurations from a Different Account**.

   Fields appear on the page. If you selected an account, it appears in the Oracle account field.

4. Type the name of the Oracle account from which to import configurations.

5. Choose a database group to which to add the configurations, and click **OK**.

   The configurations in the data account are imported to the specified database group. The configurations have statuses of invalid. For information about validating them, see Validating one or more configurations.

Copying a configuration

The copy process creates a configuration with a new name and a duplicate of the original configuration's table that stores variable attributes. The source data tables and ancillary tables associated with the original configuration are not copied. Additionally, the database group and permissions for the configuration are not copied.

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Select the row menu for a configuration (), and select **Edit**.

   The Modify Configuration page appears.

3. Click **Copy this Configuration**.

   The Add Configuration page appears.

4. Fill in the fields according to the following table, and click **Save**.

Field descriptions—Add Configuration page

Field	Description
Name	Unique name of the configuration, up to 100 characters. If a configuration was deleted by reference only, the configuration is still in the database, and you cannot add a configuration with the same name.
Description	Description that appears to end users when they browse for configurations.
Configuration Table Name	Name of the Oracle database table that will contain the variable attributes in the Oracle account to which you connected. The name must be unique in the Oracle account and must be a valid Oracle table name.

Viewing configurations

If you are a superuser, you see a list of all configurations.

If you are a non-superuser with the Manage Configurations permission, you see a list of configurations for which you have the Edit permission, either individually or from your login group.

Prerequisites

You must have the Manage Configurations permission or be a superuser.

To view configurations:

- Click **Settings**, and click **Manage Configurations**.
  The Manage Configurations page appears. All configurations appear in the table.

Editing a configuration

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Select the row menu for a configuration (), and select **Edit**.

   The Modify Configuration page appears.

3. In the last column of the first table, click the **Edit** link.

4. Select a drilldown map table, which is used for drilldown on patient counts and patient IDs, for the configuration:

   A. Next to the **Drilldown Map** field, click **Select/Edit Table**.

      The Select/Edit Drilldown Table dialog box appears.

   B. From the **Select Drilldown Table** drop-down list, select a table. The list includes all tables that are in the data account and that have the appropriate structure.

      Values for the selected drilldown table appear in the dialog box.

   C. Click **OK**.

   Note: To create or edit a table, see Creating, copying, and editing a drilldown map table.

5. Select a source description table, which is used to identify the source database tables, for the configuration:

   A. Next to the **Source Description Table** field, click **Select/Edit Table**.

      The Select/Edit Source Description Table dialog box appears.

   B. From the **Select Source Description Table** drop-down list, select a source description table. The list includes all tables that are in the data account and that have the appropriate structure.

      Values for the selected table appear in the dialog box.

   C. Click **OK**.

   Note: To create or edit a table, see Creating and editing a source description table.

6. To select an HDPS dimensions table, which displays custom dimensions in certain evaluative analysis runs:

   A. Next to the **HDPS Dimensions Table** field, click **Select Table**.

      The Select Table dialog box appears.

   B. Select a table that has been set up correctly.

C. Click **Save**.

7. Fill in the remaining fields according to the following table, and click **Save**.

The status of the configuration changes to invalid.

Field descriptions—Edit Configuration Details page

Field	Description
Name	Unique name of the configuration, up to 100 characters. The name can include the following characters: • Letters • Digits • Blank space • Plus sign (+) • Hyphen (-) • Underscore (_) • Parentheses () • Multi-byte characters • Degree symbol (°) • Registered trademark symbol (®) If a configuration was deleted by reference only, the configuration is still in the database, and you cannot add a configuration with the same name.
Description	Description that appears to end users when they browse for configurations.
Drug Hierarchy Account	Name of the Oracle account containing the hierarchy for variables with the selection type **Select From Drug Hierarchy**.
Event Hierarchy Account	Name of the Oracle account containing the hierarchy for variables with the selection type **Select From Event Hierarchy**.
Third Party Link Text	Name of the link that appears on the Patient Details page. If you do not specify text, **External Link** appears. This field is applicable only if you specify a value for **Third Party Link URL**.
Third Party Link URL	URL for the third-party application page. You can use the following substitution variables: • $REPORT_ID$ This variable is replaced by the value of the variable specified as the patient ID column for the LIST section of the drilldown map table for the configuration.

	Example: https://servername/patientinfo.asp?ptid=$REPORT_ID$ • $EXTERNAL_REPORT_ID$ This variable is replaced by the value of the external patient ID for the patient. If you use this substitution variable, the configuration must include a variable of the type **External Patient ID**.
Database Group	Database group that is used when determining color changes for patient ID links that are visited by end users. For the first radio button, select from existing database groups. The groups associated with a valid or invalid configuration in the Oracle account appear. For the second radio button, you can type the name of a database group to create.
Group Permissions Individual Permissions	See Assigning permissions to a configuration.

Validating one or more configurations

You can validate a single configuration or multiple configurations at one time. Validation is a background job that continues to run even if you log out before the process completes.

The validation process might generate warnings and errors:

- If validation generates warnings only, you can ignore the warnings so that the configuration passes validation.
- If validation produces errors, you must correct the errors before the configuration can pass validation.

For information about the checks that are performed, see Validation checks for configurations.

Steps performed during the validation process

The validation process performs the following steps:

1. Validation checks the configuration and its variables.
2. Validation creates UVTs.

   If you start validation from the Modify Configuration page by clicking **Validate Now**, UVTs are created only if no configuration errors occur.

   If you start validation from the Manage Configurations page by clicking the **Validate Configurations** link, UVTs are created even if configuration errors occur.
3. Validation checks source data.
4. If there are no data errors, validation creates Patient ID tables for any variables for which the **Patient ID Table** attribute has been set to the name of a table that does not already exist in the data account.

To validate configurations:

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.
2. To validate a single configuration:

   A. Select the row menu for a configuration (), and select **Edit**.

   The Modify Configuration page appears.

   B. Click **Validate Now**.

   The validation process runs, and one of the following results occurs:

   - The configuration is valid and has no errors or warnings.
   - The configuration is valid with warnings.

- The configuration is invalid with errors and possibly warnings.

If the validation generates warnings, review the warnings. To ignore them, select **Ignore Warnings**.

If no errors were found and you ignore the warnings, the configuration is validated.

If both warnings and errors were found and you ignore the warnings, the page reappears with the same warning messages and error messages.

3. To validate multiple configurations:

   A. Click **Validate Configurations**.

      The Configuration Validation page appears.

   B. Select the configurations to validate.

      A configuration is selected by default if it is invalid.

   C. Click **Validate**.

   The validation process runs, and one of the following results occurs:

   - The configuration is valid and has no errors or warnings.
   - The configuration is valid with warnings.
   - The configuration is invalid with errors and possibly warnings.

4. If the validation generates warnings, review the warnings and decide if they are acceptable.

Deleting a configuration

When a configuration is deleted, ancillary tables associated with the configuration remain in the database.

The options for deleting a configuration depend on whether you are a superuser or a non-superuser with the Manage Configurations permission.

To delete a configuration:

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Select the row menu for a configuration (), and select **Edit**.

   The Modify Configuration page appears.

3. Click **Delete this Configuration**.

Results for a superuser

If you are a superuser, you are prompted to choose a deletion method:

- **Remove Reference Only**—The deletion process removes the configuration from the Empirica Healthcare Analysis software but not from the source data account. If you choose this option, you can import the configuration later.
- **Delete Permanently**—The deletion process removes the configuration from the Empirica Healthcare Analysis software and from the source data account.

If the configuration has report definitions, queries, or saved lists, the software prompts you to choose one of the following options for handling the objects that reference the configuration:

- From a drop-down list, choose the configuration to which compatible objects should be transferred. For an object to be compatible with another configuration, the variables used by the object must exist in the other configuration and must have the same variable name, selection type, and Oracle data type.
- Choose to delete all objects referenced by the configuration.

If the configuration has no report definitions, queries, or saved lists, but at least one of any of the following objects, a warning message appears, informing you that the objects will be deleted if you continue:

- Analysis run
- Patient group
- Report output

Results for a non-superuser

If you are not a superuser, and if the configuration has any analysis runs, queries, patient groups, saved lists, report definitions, or report outputs, a message appears, informing you that the configuration cannot be deleted because it is referenced by objects.

Note: If this message appears and you do not see any objects associated with the configuration, the objects might not be published to you, or you might not have the necessary permissions to see the objects.

If the configuration has no analysis runs, queries, patient groups, saved lists, report definitions, or report outputs, the software prompts you to choose one of the following options for removing the configuration:

- **Remove Reference Only**—The deletion process removes the configuration from the Empirica Healthcare Analysis software but not from the source data account. If you choose this option, you can import the configuration later.
- **Delete Permanently**—The deletion process removes the configuration from the Empirica Healthcare Analysis software and from the source data account.

Working with variables

Required variables

For information about adding a variable, see Adding or editing a variable.

The following variables are required, except where noted.

Table	Type, Subtype	Number to create	Selection type	Prefix	Notes
Demographics	Patient ID	1	Free Text	None	
Demographics	Other, Enrollment Start Date	1	Continuous	None	
Demographics	Other, Enrollment End Date	1	Continuous	None	
Demographics	Gender	1	Distinct Value List	None	
Demographics	Age, Age at enroll-start or Age, Age at enroll-end	1	Continuous	None	
Demographics	Other, Age Group	1	Distinct Value List	None	Optional. You can create this variable if the source data contains this information.
Demographics	Other, None	1 for each variable	Varied	None	Optional. Use this variable to make additional demographic variables available for stratification, querying, reporting, or drilldown.
Drugs	Patient ID	1	Free Text	None	
Drugs	Drug	1 for each level of the	Select From Drug	D	These drug variables are

		hierarchy	Hierarchy		available in analysis runs.
Drugs	Record Date	1	Continuous	None	
Drugs	Record Duration	1	Continuous	None	
Drugs	Other, Drug Label	1	Distinct Value List	None	
Drugs	Other, None	1 for each variable	Varied	None	Optional. Use this variable to make additional drug variables available for querying, reporting, or drilldown.
Diagnoses	Patient ID	1	Free Text	None	
Diagnoses	Event, None	1 for each level of hierarchy	Select from Event Hierarchy	E	These event variables are available in analysis runs.
Diagnoses	Record Date	1	Continuous	None	
Diagnoses	Other, Event Label	1	Distinct Value List	None	
Diagnoses	Inpatient Y/N	1	Distinct Value List	None	Optional. Use this variable to distinguish inpatient from outpatient diagnoses in single-patient timelines.
Diagnoses	Other, None	1 for each variable	Varied	None	Optional. Use this variable to make additional

					diagnosis variables available for querying, reporting, or drilldown.
Procedures	Patient ID	1	Free Text	None	
Procedures	Event, Procedure	1	Distinct Value List	P	This procedure variable is available in analysis runs.
Procedures	Record Date	1	Continuous	None	
Procedures	Other, Proc Label	1	Distinct Value List	None	
Procedures	Other, Inpatient Y/N	1	Distinct Value List	None	
Procedures	Other, None	1 for each variable	Varied	None	Optional. Use this variable to make additional procedure variables available for querying, reporting, or drilldown.

Additional notes for variables

- To make a variable available for breakdowns in descriptive analysis runs and covariates in evaluative analysis runs, create it as a stratification variable. Typically, a stratification variable includes two or three demographic variables, such as gender, race, or age group.
- To make a variable available in Browse Data, associate a patient ID table with it. Typically, you make the following variables available in Browse Data:
  - One or more demographics variables.
  - One drug variable.
  - One diagnosis variable
  - One procedure variable, if the source data contains procedures.

- Determine which variables to show on the Patients page and the Patient Details page. The drilldown map table determines the variables that appear.

Adding or editing a variable

For a list of the required variable, see Required variables.

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Select the row menu for a configuration ( ), and select **Edit**.

   The Modify Configuration page appears.

3. To add a variable, click **Add New Variable**.

   or

   To edit a variable, select the row menu for a variable ( ), and select **Edit**.

   The Edit Variable page appears.

4. Specify a unique values table (UVT):

   Note: If the number of unique values for a variable is unusually high, selecting values from a list might not be practical. In this case, you can set up the variable with the selection type of **Free Text**.

   A. To create a new UVT during validation, leave the **Unique Values Table** field blank.

   B. To select an existing table:

      i. Next to the **Unique Values Table** field, click **Select Table**.

         The Select Table dialog box appears, displaying all tables in the data account.

      ii. Select a table with the appropriate structure, and click **Save**.

5. To specify a patient ID table:

   A. To create a new patient ID table, in the **Patient ID Table** field, type a name that does not exist in the data account.

      The patient ID table is generated during validation and associated with the variable.

   B. To specify no patient ID table, leave the **Patient ID Table** field blank.

      A patient ID table is not created. No patient ID table is used for the variable.

      If a saved variable contains a value in the field and you delete the value, the patient ID table is no longer associated with the variable. The previously specified patient ID table is not removed from the data account.

C. To select an existing table:

i. Next to the **Patient ID Table** field, click **Select Table**.

The Select Table dialog box appears, displaying all tables in the data account.

ii. Select a table with the appropriate structure, and click **Save**.

The application checks to make sure that the format of the patient ID table complies with the required table format.

6. Fill in the remaining fields according to the following table, and click **Save**.

If you edited a variable, the status of its configuration changes to invalid.

Field descriptions—Edit Variable page

Field	Description
Name	Name of the variable, also called a configuration variable, up to 100 characters. The name appears to end users when they select a variable to use. The name can include the following characters: • Letters • Digits • Blank space • Plus sign (+) • Hyphen (-) • Underscore (_) • Parentheses () • Multi-byte characters
Description	Description of the variable.
Table	Name of source database table or view containing the data that is represented by the variable. To select the table and column, click **Select Table/Column**. You can select any table in the Oracle account to which you are connected, and any column in that table.
Column	Name of the database column containing the data that is represented by the variable.
Prefix	Prefix of the variable. A drug, event, or procedure variable must have a prefix for users to be able to select it while creating an analysis run. • To make a drug variable available for analysis, you must use the prefix **D**. Typically, you want all drug variables that represent levels of the

	hierarchy to be available for analysis. • To make an event variable available for analysis, you must use the prefix **E**. Typically, you want all event variables that represent levels of the hierarchy to be available for analysis. • To make a procedure variable available for analysis, use the prefix **P**, the **Event** type, and the **Procedure** subtype. Create only a single procedure variable this way. Do not include a prefix for any other variables or for variables with the following subtypes: • Drug Label • Event Label • Proc Label
Variable Type	See Variable types.
Variable Subtype	(Appears only for certain variable types.) See Variable subtypes.
Use as stratification variable	When selected, the variable can be used for breakdowns of descriptive analysis or covariates in evaluative analysis. Select this option only for demographic variables.
Selection Type	See Variable selection types.
Hierarchy Level	Level of the hierarchy represented by the variable. This field is editable if the selection type of the variable is **Select From Drug Hierarchy** or **Select From Event Hierarchy**. Otherwise, the field is read-only, and N/A is selected.
Hide From Query Wizard	• **Yes**—The variable is not available in the Query Wizard. On the Modify Configuration page, **Analyses** appears in the Visibility column for the variable. • **No**—The variable is available in the Query Wizard. On the Modify Configuration page, **Analyses & Queries** appears in the Visibility column for the variable.

Deleting a variable from a configuration

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Select the row menu for a configuration (), and select **Edit**.

   The Modify Configuration page appears.

3. In the second table, select the row menu for the variable, and select **Delete**.

4. A confirmation message appears.

5. Click **OK**.

   If no runs reference the variable, the variable is deleted. Any unique values table or patient ID table that is referenced by the variable remains in the database.

   If a run references the variable, a message appears, informing you that the variable cannot be deleted. The message appears if the variable is used in any of the following ways:

   - An index event or outcome event in a descriptive or evaluative analysis run.
   - A co-occurrence window or history window in a descriptive or evaluative analysis run.
   - A selected breakdown or covariate in a descriptive or evaluative analysis run.
   - A variable in a patient statistics run.

Unique values tables (UVTs)

The unique values table (UVT) is a one-column table that lists all unique values for a single variable from the source data. UVTs increase the efficiency of value retrieval for lists of available values. In many places in the Empirica Healthcare Analysis application, users select from a list of available values for a particular variable. Typically, the list of available values for a variable is constructed from the unique values table for the variable.

When creating or editing a configuration, you have the following options:

- Associate existing UVTs with appropriate variables.

  You can specify a UVT for variables with the following selection types:

  - Distinct Value List
  - Select From Drug Hierarchy
  - Select From Event Hierarchy
- Request that the software creates the UVTs during validation of the configuration.

  If you do not specify a UVT for a variable, the validation process creates the UVT.

UVTs generated during validation

The validation process generates a UVT for any variable that meets all the following criteria:

- The variable has one of the following selection types:
  - Distinct Value List
  - Select from Drug Hierarchy
  - Select from Event Hierarchy
- The variable has no unique values table name already specified.

Enrollment dates

Automatically generated UVTs include all values that occur in the data, regardless of whether patients have enrollment start and stop dates and regardless of whether values occur within enrollment periods.

Patient ID tables

A patient ID table is required for each variable you want available in the Browse Data feature. If no variables in a configuration have an associated patient ID table, the configuration does not appear in Browse Data.

Typically, two or three demographic variables, such as gender, age group, or race, appear in the Browse Data dialog box. Additionally, you should include only one drug variable, one event variable, and one procedure variable.

Note: Age Group is available in the Browse Data dialog box only if an age group variable exists in the source data, and the variable is associated with a Patient ID table. The Browse Data feature does not require the **Age Group** variable subtype.

You have the following options for creating patient ID tables:

- Create the patient ID tables outside the Empirica Healthcare Analysis application, and then associate the tables with the appropriate variables in the application. The tables must be in the data account.
- Generate the patient ID tables during the validation of the configuration.

The Patient ID Table attribute is available for only variables with the following selection types:

- Distinct Value List.
- Select from Drug Hierarchy.
- Select from Event Hierarchy.

A generated table contains one row for each unique combination of the variable value and patient identifier that occurs in the source table containing the variable.

An automatically created patient ID table includes the following:

- For a demographic variable, only patients who have both an enrollment start date and an enrollment end date.
- For event or procedure variables, only events or procedures that occur within a patient's enrollment period. The patients must have enrollment start and end dates.
- For drug variables, only prescriptions that start but do not necessarily end within a patient's enrollment period. The patients must have enrollment start and end dates.

Note: If you create a patient ID table outside the application, make sure to include only the patients, events, procedures, and prescriptions that meet these rules.

Variable types

A variable maps to a column in the source data and allows you to use the source data in analysis runs. The type and subtype of a variable determine where and how the source data is available in the application.

Variable type	Description	Additional requirements
Age	A variable with this type allows you to use the computed age for breakdowns in analysis runs. For information about the available subtypes for this type, see Variable subtypes.	The selection type must be one of the following options: • Distinct Value List • Continuous
Patient ID	A variable with this type identifies a patient ID. You must include a variable with this type for each source database table referenced by a variable. As a result, the configuration might have several Patient ID variables. However, you can include only one variable of this type for each source database table.	The selection type must be Distinct Value List.
Drug	A variable of this type identifies a drug. To make a drug variable available in analysis runs, create the variable using this type and include the prefix **D**.	
Event	A variable of this type identifies an event (diagnosis) or procedure. To make an event or procedure variable available in analysis runs, use this type when creating the variable, and include the prefix **E** for a diagnosis or **P** for a procedure. For information about the available subtypes for this type, see Variable subtypes.	You must use the following subtypes: • An event variable must have no subtype. • A procedure variable must have the **Procedure** subtype.
External Patient ID	Required to include a link to a third-party application from patient details. You must also specify an external application URL in the Third Party Link URL when you edit a configuration. To use this feature, there must be a mapping of Empirica Healthcare Analysis patient IDs to the IDs used by the external application. For example, in the demographics table, you	

	could include an additional column named EXTERNAL_ID, and populate it with the external ID associated with each patient ID. In the configuration, you must include a variable for the column with the External Patient ID variable type. If the patient IDs used by the external application are the same as the patient IDs in the application, you do not need an external patient ID to link to the external application.	
Gender	A variable of this type allows you to use the gender variable for breakdowns or covariates in analysis runs, as long as the variable is marked as a stratification variable.	The selection type must be one of the following options: • Distinct Value List • Free Text
Generalized Item	Reserved for future use.	
Other	A variable of this type identifies additional required information and is classified by the subtype of the variable. The column in the source data must have an Oracle data type of CHAR, VARCHAR2, NUMBER, FLOAT, INTEGER, DECIMAL, or SMALLINT. For information about the available subtypes for this type, see Variable subtypes.	
Record Date	A variable of this type represents the following: • In a drug table, the start date of a prescription. • In an event table, the occurrence date of an event. • In a procedure table, the occurrence date of a procedure. Create only one variable with this type for each source database table.	The selection type must be Continuous.
Record Duration	A variable of this type represents the number of days supplied for a prescription. Create this variable only for a drug table.	The selection type must be Continuous.

Variable subtypes

Some variable types require you to specify a subtype. If a variable has a subtype, the subtype appears in parentheses after the variable type on the Modify Configuration page.

Variable subtypes for the Age variable type

Variable subtype	Description
None	A variable with this subtype represents data that does not fit any of the following categories. Computed Age is not available for breakdowns or covariates in analysis for Age variables of this subtype.
Age at enroll-start	This subtype identifies the age variable as being the patient's age at the start of that patient's enrollment. This variable subtype allows you to use computed age or computed age group for breakdowns or covariates in analysis runs. You do not need to mark a variable using this subtype as a stratification variable.
Age at enroll-end	This subtype identifies the age variable as being the patient's age at the end of that patient's enrollment. This variable subtype allows you to use computed age or computed age group for breakdowns or covariates in analysis runs. You do not need to mark a variable using this subtype as a stratification variable.

Variable subtypes for the Event variable type

Variable subtype	Description
None	A variable with this subtype represents an event.
Procedure	A variable with this subtype represents a procedure. Create only one procedure variable, and use the **P** prefix for it.

Variable subtypes for the Other variable type

Variable subtype	Description
None	A variable with this subtype represents data that does not fit any of the following categories.
Enrollment	A variable with this subtype represents the start date of enrollment.

Start Date	
Enrollment End Date	A variable with this subtype represents the end date of enrollment.
Age Group	A variable with this subtype represents an age group in the source data. You can use the age group for breakdowns in analysis runs, if the variable is marked as a stratification variable.
Drug Label	A variable with this subtype represents the drug variable that appears by default the first time you view single-patient timelines.
Event Label	A variable with this subtype represents the event variable that appears by default the first time you view single-patient timelines.
Proc Label	A variable with this subtype represents the procedure variable that appears by default the first time you view single-patient timelines.
Inpatient Y/N	A variable with this subtype indicates whether an event or procedure is inpatient or outpatient. If specified, inpatient and outpatient events or procedures are differentiated in single-patient timelines.

Variable selection types

Selection type	Description
Distinct Value List	End users choose values from a list. If **Use as Stratification Variable** is selected for a variable, you must use this selection type for the variable.
Select From Drug Hierarchy	End users select from a hierarchy of values. You must specify the following: • The **Drug Hierarchy Account** attribute of the configuration. • The **Hierarchy Level** attribute of the variable.
Select From Event Hierarchy	End users select from a hierarchy of values. You must specify the following information: • The **Event Hierarchy Account** attribute of the configuration. • The **Hierarchy Level** attribute of the variable.
Continuous	End users specify a range of values by entering From and To values. Use this selection type only for numeric or date time columns in the source data.
Free Text	End users type a text string for which matches are found.

Ancillary tables for configurations

Ancillary tables

When setting up a configuration, you must refer to the following ancillary tables, in addition to the tables containing the source data.

Prerequisites and sharing

Ancillary tables must be in the data account.

Because multiple configurations can exist for the same data account, multiple configurations can refer to the same ancillary table. When modifying an ancillary table, remember that other configurations for that data account can be affected.

For example, suppose that a drilldown map table named DRILLDOWN_MAP is shared by multiple configurations. If you edit a configuration and change DRILLDOWN_MAP to show drug route on the Patient Details page, all configurations in the account that use DRILLDOWN_MAP now have drug route on the Patient Details pages.

Structure and content

All ancillary tables except the HDPS dimensions table can be created from within the Empirica Healthcare Analysis application. A table created from within the application always has the appropriate structure. If you create an ancillary table from the back end, make sure the table has the appropriate structure and is not empty.

Tables

Table name	Purpose	Creation and reference
Drilldown map table	Populates data when you drill down to lists of patients and patient details.	You can create, copy, and edit this table in the application. You can also create this table outside the application. This table is referenced by a configuration.
Source description table	Identifies database tables containing source data.	You can create and edit this table in the application. You can also create this table outside the application. This table is referenced by a configuration.
HDPS Dimensions table	Required only if end users use custom HDPS dimensions. Used by evaluative analysis	You create this table outside the application. This table is referenced by a

	runs that use high-dimensional propensity scores.	configuration.
Unique values table	Populates lists of terms from which users select values.	This table is generated when you validate a configuration. You can also create this table outside the application. This table is referenced by a variable.
Patient ID table	Required by the Browse Data feature.	This table is generated when you validate a configuration. You can also create this table outside the application. This table is referenced by a variable.

Deletion

You cannot delete ancillary tables from within the Empirica Healthcare Analysis application.

Creating, copying, and editing a drilldown map table

The drilldown map table that is associated with a configuration determines the type of data presented during drilldown, which consists of the following activities:

- For a patient count in analysis results or elsewhere, you click **View Patients** for a patient count.
- For the Patients page, you click patient ID links to drill down to a list of patients.

After you edit a drilldown map table, you must validate any configurations that use the table.

To create, copy, or edit a drilldown map table:

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Select the row menu for a configuration (), and select **Edit**.

   The Modify Configuration page appears.

3. In the last column of the first table, click the **Edit** link.

   The Edit Configuration Details page appears.

4. Next to the **Drilldown Map** field, click **Select/Edit Table**.

   The Select/Edit Drilldown Table dialog box appears.

5. To create a new table:

   A. Click **Create New Table**.

      The content of the dialog box changes.

   B. Type the name of the table, and click **Create Table**.

      A table with the following types appears in the dialog box.

Type	Description
LIST	Determines columns of the Patients page. The first variable represents the patient ID.
CASE	Determines columns of the Patient Information section of the Patient Details page. The patient ID variable is included automatically as the first column in the section.
DRUG	Determines columns of the Reported Drugs/Vaccine section of

	Patient Details page. The patient ID variable is included automatically as the first column in the section.
EVENT	Determines columns of the Reported Events/Symptoms section of the Patient Details page. The patient ID variable is included automatically as the first column in the section.
PROCEDURES	Determines columns of the Procedures section of the Patient Details page. The patient ID variable is included automatically as the first column in the section.
NARRATIVE	Narrative text, such as medical history for the patient, which appears on the Patient Details page. If multiple rows per patient ID exist, select the **NARRATIVE spans multiple rows** option.

Note: When you add or edit a configuration, an error message appears if the drilldown map table does not have LIST specified. During validation, warning messages appear if the drilldown map table does not have CASE, DRUG, or EVENT specified.

C. Next to each of the types described in the previous table, click **Edit**.

Fields appear in the Select/Edit Drilldown Table dialog box.

D. Fill in the fields according to the following table, and click **Save**.

6. To copy a table:

A. From the **Select Drilldown Table** drop-down list, select a table, and click **Copy Table**.

A dialog box appears.

B. Type the name of the table, and click **Create Table**.

7. To edit a table:

A. From the **Select Drilldown Table** drop-down list, select a table to edit.

B. In the table, find the type to edit, and click the **Edit** link for the row. For a description of each of the types, see the previous table.

Fields appear in the Select/Edit Drilldown Table dialog box.

C. Fill in the fields according to the following table, and click **Save**.

D. To clear the specification for an information type, click **Clear** next to the specification.

8. Click **OK**.

Select/Edit Drilldown Table dialog box—Field descriptions

In this dialog box, you specify the columns that appear in drilldown.

Field	Description
<type> Table Patient ID Col	Click the **Select Patient ID Field** button and select values for the following: • The source data table. • The column of that table that represents the patient ID.
Order By	Type the name of a source data table column that is used to sort rows alphanumerically for the information type. For example, you might want to sort the Patients page, determined by the LIST section. Separate multiple columns by a comma.
Section Title	Type a label that appears as the section heading.
Column	Select the source table column for which information is displayed for the information type. The dropdown list includes all columns in the source data table. Columns appear in drilldown in the order in which you specify them. For a currently specified column, you can select a **—delete—** entry. You can select a column only one time for a given information type.
Label	Type a label, up to 30 characters, that appears as the column heading. If you do not specify a value, the column name, which appears in the **Column** field, is used for the column heading. The value cannot contain a question mark.
Distinct	When selected, only distinct values for the column appear in drilldown.

Creating and editing a source description table

A source description table identifies the tables containing source data for the configuration.

You should create a separate source description table for each configuration that references different source database tables. For example, if two configurations in the same Oracle account reference different database tables, each configuration should have its own source description table.

After you edit a source description table, you must validate any configurations that use the table.

1. Click **Settings**, and click **Manage Configurations**.

   The Manage Configurations page appears.

2. Select the row menu for a configuration (), and select **Edit**.

   The Modify Configuration page appears.

3. In the last column of the first table, click the **Edit** link.

   The Edit Configuration Details page appears.

4. Next to the Source Description Table field, click **Select/Edit Table**.

   The Select/Edit Source Description Table dialog box appears.

5. To create a table:

   A. Click the **Create New Table** link.

   B. Type a name for the dialog box, and click **Create Table**.

      The table is created.

   C. Fill in the fields according to the following dialog box, and click **Save**.

6. To edit a table:

   A. From the **Select Source Description Table** drop-down list, select a table.

      Values appear in the fields in the dialog box.

   B. Fill in the fields according to the following table, and click **Save**.

Field descriptions—Select/Edit Source Description Table

Field	Description
Select Source Description Table	Source description table that is associated with the configuration. The drop-down list contains all tables in the data account with the appropriate structure.

Description	Description of the data source. This value appears at the bottom of Patient Details pages.
Load Date	Date the source data was loaded. This value does not appear anywhere else in the user interface.
Data Source	Source of the data. This value does not appear anywhere else in the user interface.
Demographics Table	Name of the demographics table for which variables are defined. You can type a table name or click the **Select Table** button to select an existing table. If you do not provide a value, an error occurs during validation.
Drug Table	Name of the drug table for which variables are defined. You can type a table name or click the **Select Table** button to select an existing table. If you do not provide a value, an error occurs during validation.
Event Table	Name of the event table for which variables are defined. You can type a table name or click the **Select Table** button to select an existing table. If you do not provide a value, an error occurs during validation.
Procedure Table	Name of the procedure table for which variables are defined. You can type a table name or click the **Select Table** button to select an existing table.

Custom HDPS dimensions table

A custom HDPS dimension table allows the creator of an evaluative analysis run to select one or more custom dimensions, such as inpatient diagnoses or non-maintenance drugs. The selected dimensions are evaluated for their possible influence on the propensity of patients to experience the primary index event.

To select custom dimensions, you must set up an HDPS dimensions table outside the Empirica Healthcare Analysis software and then associate the table with a configuration. The table must be in the data account. Each configuration in an account can have a different HDPS dimensions table.

The run wizard for an evaluative HDPS run checks whether a custom HDPS dimensions table is specified as an attribute of the configuration:

- If such a table exists, the table determines the fields to show as available dimensions on the Select HDPS Dimensions page.
- If a HDPS dimensions table is not specified, or if the specified table does not meet the expected structure of the dimensions table, standard dimensions are available instead of custom dimensions. The standard dimensions are for each variable where the variable type is drug, event, or procedure, and a patient ID table is specified.

Table contents

The HDPS dimensions table must contain one row for each custom dimension to appear on the HDPS dimensions page in the evaluative analysis run wizard. The table must have the following columns.

Column	Structure	Content
DIM_NAME	VARCHAR2 (100 CHAR) not null	Label that appears on the HDPS dimensions page.
DIM_VARNAME	VARCHAR2 (100 CHAR) not null	Name of the variable. This value is the variable name, not the column name.
DIM_WHERE	VARCHAR2 (2000 CHAR))	WHERE clause to restrict values of the variable. This field uses the column name, not the variable name. If you do not want to restrict values, leave this column null.

To support both standard dimensions and custom dimensions, you can set up the dimensions table with rows that have an additional restriction and rows that do not. For example, the following table supports both standard and custom variables for diagnoses and drugs.

DIM_NAME	DIM_VARNAME	DIM_WHERE
Any ICD9 code plus text (5-digit level)	ICD9 code plus text (5-digit level)	
Inpatient ICD9 code plus text (5-digit level)	ICD9 code plus text (5-digit level)	INPATIENT='Y'
Generic name	Generic name	
Generic name for non-maintenance drugs	Generic name	MAINTENANCE='N'

Patient statistics runs

Analysis Setup page

On the Analysis Setup page, you can:

- View information about existing patient statistics runs.
- Create patient statistics runs.

General activities

The following links and filters appear at the top of the page and affect the entire page.

- Create Patient Statistics Run
- Columns
- Print
- Download
- Select Rows
- Filter by Project or Configuration

Row-specific activities

The following menu options are available from the row menu, located in the leftmost column of the table, and affect an individual row in the table. Your permissions determine the options that appear.

- View Results
- Rename
- Cancel (for runs that are in progress) or Delete (for completed runs)
- Re-run
- Create Definition File
- View Jobs for Run
- View Run Details
- Publish

Field descriptions—Analysis Setup page

Field	Description
Configuration	Name of the data configuration on which the patient statistics run is based.
Created	Date and time the patient statistics run was created.
Created By	User name of the superuser who created the patient statistics run.
Description	Description of the patient statistics run.

End Date	Date and time the patient statistics run ended. This column is empty until all component jobs are complete.
ID	Identifier that was assigned automatically when the patient statistics run was submitted. Run IDs are unique across the Analysis Setup, Descriptive Analysis, and Evaluative Analysis tabs. IDs are not reused when a run is deleted.
Name	Name of the patient statistics run.
Project	Name of the project to which the patient statistics run is assigned.
Run Type	**Patient Stats** appears.
Size	Size of the complete results table.
Start Date	Date and time the run started. For a scheduled run, the start date might be substantially later than the creation date of the run.
Status	One of the following values: • **Completed**—All component jobs completed successfully, and the run is complete. • **Error Occurred**—A job for the run failed, so the run failed. • **Cancelled**—The run was cancelled.

Patient statistics runs

A user with the Manage Configuration permission can create a patient statistics run and view its results.

A patient statistics run supports the following activities:

- When you click **View Preliminary Counts** while creating a descriptive or evaluative analysis run, the counts are produced using the results of a patient statistics run.
- In an evaluative analysis run, the results of a patient statistics run are used for the screening process by which patients with both the primary index event and the comparator index event are dropped from the analysis.

A patient statistics run should be executed for a configuration before anyone uses the configuration to perform descriptive or evaluative analysis.

Variables used in the run

When you create a patient statistic run, you do not select variables. Variables selection is automatic and includes:

- The drug variable that is at the lowest level of the hierarchy and that has a prefix.
- The event variable that is at the lowest level of the hierarchy and that has a prefix.
- The procedure variable with a prefix.

Enrollment dates

A patient statistics run counts only drugs, diagnoses, and procedures that occur within patient enrollment periods. For drugs, the prescription must start, though not necessarily end, within patient enrollment periods. The run looks at prescriptions, not drug eras.

A patient statistics run does not include drugs, diagnoses, or procedures for a patient who is missing either an enrollment start, enrollment end date, or both.

Usage

A descriptive analysis or evaluative analysis run uses the most recently executed patient statistics run that meets all the following criteria:

- The patient statistics run is for the same configuration as the analysis run.
- The patient statistics run completed successfully.
- The patient statistics run has been published to a login group, and the user creating the descriptive or evaluative analysis run is a member of the login group.

Creating a patient statistics run

A patient statistics run is a background job. After the run completes, you can view the results of the run.

Prerequisites

- The configuration for which you want to create a patient statistics run must be valid.
- You must have the Manage Configurations permission.
- A drug hierarchy must be set up and associated with drug variables in the configuration.
- An event hierarchy must be set up and associated with event variables in the configuration.
- The variable representing the lowest level of the drug hierarchy must have a variable prefix.
- The variable representing the lowest level of the event hierarchy must have a variable prefix.
- The procedure variable to be used in analysis runs must have a variable prefix.

To create a patient statistics run:

1. Select the **Analysis Setup** tab.
2. Click **Create Patient Statistics Run**.

   The Create Patient Statistics Run page appears.
3. From the drop-down list, select the data configuration for the patient statistics run to use, or click **Browse** to navigate to a configuration.
4. Click **Next**.

   The Patient Statistics Run: Advanced Options page appears.
5. In the **Maximum memory available to metrics-generation engine** field, type a value for the amount of memory to devote to the engine.

   The value that appears is the lesser of the estimate of the memory required for the run and the maximum allowed value. The estimate varies from run to run.

   The remaining information on the page is read-only.
6. Click **Next**.

   The Run Options page appears.
7. Fill in the fields, and click **Next**.

   The Name Analysis Run page appears.
8. Fill in the fields, and click **Next**.

The Confirm Run Parameters page appears.

9. Review the settings, and click **Submit Run**.

   After the run is submitted, the Run Submitted page appears.

10. Click **Continue**.

    The Analysis Setup page appears. For information about the status of the run, check the Status column.

    For information about viewing the results of the run, see Viewing the results of a patient statistics run.

Viewing the results of a patient statistics run

The results of a patient statistics run appear in a tabular format.

Run results include a row for each unique value of the following:

- The drug variable that represents the lowest level in the event hierarchy and has a variable prefix.
- The event variable that represents the lowest level in the drug hierarchy and has a variable prefix.
- The procedure variable that has a variable prefix.

To view the results of a patient statistics run:

1. Select the **Analysis Setup** tab.
2. To filter the patient statistics runs that appear, select options from the **Project** and **Configuration** drop-down lists.
3. Navigate to the row containing the run for which to view results.

   Note: The run must have completed successfully. To verify that the run has completed, check the **Status** column.

4. Select the row menu (), and select **View Results**.

   The Run Results page appears.

Field descriptions—Run Results page

If a count is a link, you can click the link to drill down.

Field	Description
Item	Drug, diagnosis, or procedure.
PATDRG1	For drugs, the count of patients who have the value in the Item column.
PATEVT	For diagnoses or procedures, the count of patients who have the value in the Item column.
Prefix	One-letter prefix of the variable for which the Item column shows values. The prefix is an attribute of a variable in a configuration.

Publishing a patient statistics run

You must publish a patient statistics run to one or more login groups for the run to be available for use by descriptive analysis and evaluative analysis runs.

You can publish to your login group. If you are a superuser, you can publish to any login group.

1. Select the **Analysis Setup** tab.
2. To filter the patient statistics runs that appear, select options from the **Project** and **Configuration** drop-down lists.
3. Navigate to the row containing the run to publish.
4. Select the row menu (), and select **Publish**.

   The Publish Run page appears.
5. Select the login groups to which to publish the run. Use the **Shift** and **Ctrl** keys to select multiple login groups.
6. Click **Publish**.

   The run is published.
7. To return to the Analysis Setup page, click the **Back** link at the bottom of the page.

Creating and running a definition file

For an existing analysis run, you can create a definition file that allows you to submit the analysis run automatically instead of through the Empirica Healthcare Analysis user interface. You can run the definition file as often as needed.

Prerequisites

- You must be a superuser.
- The analysis run must have completed.

To create and run a definition file:

1. Select one of the following tabs:
   - Descriptive Analysis
   - Evaluative Analysis
   - Analysis Setup
2. Select the row menu () for an analysis run, and select **Create Definition File**.

   The Run Definition File dialog box appears and contains the contents of the file.
3. Select the full text in the dialog box.
4. Paste the text into a text application, such as Notepad.
5. To modify the name of the analysis run, edit the **name** parameter.

   Note: Do not modify the values of any other parameters.
6. Save the document with an extension of **IN**, and close the document.
7. Copy the file to the correct directory:
   A. Open the webvdme.properties file, located in the following directory on the application server:

      $INSTALL_DIR/Healthcare/WEB-INF/classes

   B. Find the temp_dir parameter, and note the directory it references, such as:

      temp_dir=/u01/app/oracle/product/EHC/temp

   C. Navigate to the specified directory, and locate the **input** folder.

      For example, for the previous example, the directory is

      /u01/app/oracle/product/EHC/temp/input

   D. Place the IN file in the input directory.

The listener process checks the input directory every 2 to 30 seconds and executes any run definition files with the IN extension. While being processed, the file has an extension of PROC. The file extension changes to DONE when the execution of the run definition file is complete.

Common tools

Analysis run information

Determination of the index date

The way the index date is determined depends on the following information:

- The variable type that was used for the index event.
- The type of index event.

Option descriptions—Define index event as _____ of terms selected below

Option	Description
For a diagnosis or procedure variable: Start of first occurrence	For each patient, the analysis identifies an index date candidate by looking at the first occurrence of the index event during both the patient's enrollment period and the analysis period. The analysis applies user-specified temporal parameter criteria to determine the follow-up for the patient, with one of the following results: • If the index lead-in and outcome lead-in criteria would be met, the index date candidate becomes the index date for the patient. • If the index lead-in and outcome lead-in criteria would not be met, the analysis looks at the next occurrence of the diagnosis or procedure event during both the patient's enrollment period and the analysis period and applies the temporal parameter criteria again. This process is repeated until the temporal criteria are met. • If no index date candidate meets the index lead-in and outcome lead-in criteria, an index date is not set for the patient.
For a drug variable: Start of first occurrence	For each patient, if no index lead-in and no outcome lead-in are specified for the analysis, the analysis identifies an index date candidate as the first day of a drug era where that day is during both the patient's enrollment period and the analysis period. The drug era must start within the enrollment period but does not necessarily need to start within the analysis period. The drug era does not need to end within the enrollment period or analysis period. If no such drug era exists, an index date is not set for the patient. **Note:** When lead-ins are not specified, if a drug era starts before the analysis period and continues into it, the index date candidate is the same as the analysis period start date. For each patient, if an index lead-in of 1 or more or an outcome lead-in of 1 or more is specified for the analysis, the analysis identifies an index date candidate as the first day of the first

	drug era that starts during the patient's enrollment period and the specified analysis period. The drug era does not need to end within the enrollment period or analysis period. The analysis applies user-specified temporal parameter criteria to determine the follow-up for the patient, with one of the following results: • If the index lead-in and outcome lead-in criteria would be met, the index date candidate becomes the index date for the patient. • If the index lead-in and outcome lead-in criteria would not be met, the analysis looks at the first day of the next drug era that started during both the patient's enrollment period and the analysis period and applies the temporal parameter criteria again. This process is repeated until the temporal criteria are met. • If no index date candidate meets the index lead-in and outcome lead-in criteria, an index date is not set for the patient.
For a diagnosis or procedure variable: Start of last occurrence	For each patient, the analysis identifies an index date candidate by looking at the last occurrence of the index event during the patient's enrollment period and analysis period. The analysis applies user-specified temporal parameter criteria to determine the follow-up for the patient, with one of the following results: • If the index lead-in and outcome lead-in criteria would be met, the index date candidate becomes the index date for the patient. • If the index lead-in and outcome lead-in criteria would not be met, an index date is not set for the patient.
For a drug variable: Start of last occurrence	For each patient, the analysis identifies an index date candidate by looking at the first day of the last prescription that started during both the patient's enrollment period and the analysis period. The prescription does not need to end within the enrollment period or analysis period. The analysis applies user-specified temporal parameter criteria to determine the follow-up for the patient, with one of the following results: • If the outcome lead-in criteria would be met, the index date candidate becomes the index date. • If the outcome lead-in criteria would not be met, an index date is not set for the patient. When determining an index date, this option does not use drug eras, so the analysis does not combine overlapping or contiguous prescriptions and it does not use the exposure extension. **Note:** In this situation, the exposure extension is used if you define follow-up or a co-occurrence window to be the entire index exposure.

For a drug variable: End of last occurrence	For each patient, the analysis looks at all prescriptions that started during both the patient's enrollment period and the analysis period and identifies the prescription with the latest end date. The index date candidate is the last day of the prescription. If the candidate is after the patient's enrollment end or the analysis period end, an index date is not set for the patient. If the candidate is not after the enrollment end or analysis period end, the analysis applies user-specified temporal parameter criteria to determine the follow-up for the patient, with one of the following results:: • If the outcome lead-in criteria would be met, the index date candidate becomes the index date. • If the outcome lead-in criteria would not be met, an index date is not set for the patient. When determining an index date, this option does not use drug eras, so the analysis does not combine overlapping or contiguous prescriptions, and the option to specify an exposure extension is not available.
Other reasons for rejection of index date candidate	In addition to lead-in criteria not being met, the following situations result in the rejection of an index date candidate: • If the specified follow-up start would be outside of the patient's enrollment period or the analysis period: The index candidate is rejected. • If the index event is a drug: If follow-up start is specified as a number of days after index and follow-up end is specified as **end of index exposure period**, an index date candidate is rejected if the follow-up start would be later than the end of the index exposure. • If the outcome event is a drug: Even if an outcome lead-in is not specified, an index date candidate is rejected if a drug era for the drug starts before what would be follow-up start and continues into follow-up. In this situation, if the index is based on the start of the first occurrence, the analysis continues trying to find an index date. If the index is based on the start or end of the last occurrence, an index date is not set. In evaluating this condition, the analysis looks at drug eras for the outcome.

Selecting values

You select values at various steps when creating analysis runs and performing other activities.

About selecting hierarchy terms

When you select hierarchy terms, you can use a dictionary, thesaurus, or other standardized terminology that organizes or classifies terms for drugs, diagnoses, or procedures.

For example, diagnosis terms might be associated with a version of the ICD-9-CM, the International Classification of Diseases, Ninth Revision, Clinical Modification. The terms and their relationships are stored in a special account, and are associated with drug, diagnosis, or procedure variables in data configurations. If a hierarchy is associated with a variable, you can specify values for that variable by browsing that hierarchy and selecting terms.

Options for selecting values

Type values in the list box:

- Separate values with a comma. Surround an entry that contains a comma with double quotation marks, even if you include only a single value in the list.

Select values in a hierarchy account that is associated with the variable in the configuration:

1. Click **Select <*hierarchy name*> Terms**.

   The Hierarchy dialog box appears. The hierarchy section on the left lists the terms at the highest, most general level of the hierarchy.

2. In the Hierarchy list on the left, expand the tree to show terms at various levels. The hierarchy is expandable until you reach the lowest, most specific level.

   Note: Though terms at every level of the hierarchy are displayed, you can only select those terms located at or above the hierarchy level with which the currently selected variable is associated.

3. Select a term. The value appears in the **Available <*hierarchy name*> Values** list.

   or

   Select a higher-level category, and all values within the category appear in the **Available <*hierarchy name*> Values** list.

   or

   Type a value in the **Match** field (toward the bottom of the dialog box), and click **Search**. In the Hierarchy list, values that contain a match are italicized and colored blue, while values without a match are in plain, gray text. When you select a level in the hierarchy list, the Available <*level*> Values list in the middle shows all terms that match your text.

Note: When you select a term in the Available Values list (in the middle), the primary path to the term is expanded, and a **P** appears before the hierarchy terms that are in the primary path. If the term has a secondary path in the hierarchy, the path is also expanded, and an **S** appears before the terms in the path.

4. Use the arrow buttons to move the values to use in the run to the **Selected *<hierarchy name>* Values** list.

5. To save the selected values as a saved list, click **Save As List**, located below the Selected <*hierarchy name*> Values list.

6. When you are finished, click **OK**.

Select values from a drop-down list of unique values for all patients in the source data:

1. Click Select Available Values.

   A dialog box for selecting terms appears.

2. To view or change filter options, click **Show Filter Options**, and make changes as necessary. You can enable a case-sensitive search, specify the part of the values to which the filter should be applied, and select other options.

3. To filter the list, type a value in the **Filter** field, located below the Available Values list.

4. Use the arrow buttons to move values from the **Available Values** list to the **Selected Values** list.

5. To save the selected values as a saved list, click **Save As List**, located below the Selected Values list.

6. When you are finished, click **OK**.

Select a saved list of values:

1. Click **Select Saved List**.

2. Double-click a list.

   or

   Select a list, and click **OK**.

   The selected values appear in the list box.

Working with age groups

You can create custom age groups when you create an analysis run.

Rules for age groups

- The upper limit of the first age group must be greater than 0.
- After the first upper limit, each subsequent upper limit must be at least 2 more than the previous upper limit.
- Lower limits of each age group are set to a value that is one greater than the upper limit specified for the previous age group. This behavior ensures that there are no gaps between age groups.

To edit age groups:

1. Select the **Custom Age Group (of computed Age)** checkbox.

   The Custom Age Group text becomes a link.

2. Click the **Custom Age Group (of computed Age)** link.

   The Edit Custom Age Groups dialog box appears.

3. To change the values of the upper limits, modify the values.

   When you change an upper limit, the lower limit of the next group adjusts to be 1 greater than the upper limit for the previous age group.

   For example, if you have two age groups, 0-9 and 10-19, and you change the upper limit of the first group to 12, the lower limit of the next group is automatically changed to 13 when you leave the field.

4. To insert an age group:

   A. Place your cursor in the Upper Limit field for the age group after which to add a new age group.

   B. Click **Insert**.

   A new row appears. The lower limit is the upper limit of the previous group plus 1. For example, consider the following age groups:

   30 - 39

   40 - 49

   If you insert an age group after 30-39, the age groups are as follows:

   30 - 39

   40 -

   41 - 49

You must provide an upper limit for the new age group.

5. To delete an age group:

   A. Place your cursor in the Upper Limit field for an age group.

   B. Click **Delete**.

      The age group is deleted.

      If you have two age groups, 30-39 and 40-49, and you delete the 30-39 group, the remaining age group is modified to be 30-49.

      The custom age groups you define apply only to the run in which they are defined and to re-runs of the run. Runs you subsequently define use the default custom age groups unless you modify them.

How age is computed

The way that age is computed depends on whether the age variable has been configured to represent age at enrollment start or age at enrollment end.

The age computations shown below include subtractions. The number of years between two dates is computed as the number of complete years between the two dates, inclusive of both dates. For example, December 31, 2008 minus January 1, 2008 is not a complete year, whereas January 1, 2009 minus January 1, 2008 is a complete year.

How age is configured	Index	Age computation	Example
Enrollment start	Patient has an index	age in source + (index date – enroll start)	• Enrollment start = January 1, 2002 • Age = 20 • Analysis start = January 1, 2003 • Enrollment end = January 1, 2006 • Analysis end = January 1, 2007 If the patient has an index date of January 1, 2005, the computed age is 23. If the patient has no index date, the computed age is 21.
Enrollment start	Patient does not have an index	age in source + (analysis start – enroll start)	
Enrollment end	Patient has an index	age in source + (index date – enroll end)	• Enrollment start = January 1, 2002 • Analysis start = January 1, 2003 • Enrollment end = January 1, 2006 • Age = 20 • Analysis end = January 1, 2007 If the patient has an index date of January 1, 2005, the computed age is 19.
Enrollment end	Patient does not have an index	age in source + (analysis start – enroll end)	

	If the patient has no index date, the computed age is 17.

Drilling down

About drilldown

On many pages, the count of patients is a link that opens a menu with drilldown options. This menu contains the following options for patients included in the count. On some pages, additional options are available.

Option	Description
View Patients	View a list of patients. **Note:** If a count includes patients in the primary cohort and the comparator cohort, the **View Patients in Primary Cohort** and **View Patients in Comparator Cohort** menu options appear.
Single-patient Timelines	For each patient in the list, view a graph of data across time in single-patient timelines.
Multi-patient Timelines	(Available only from the results of analysis runs.) View a graph of data across time in multi-patient timelines.
Create Patient Group	Create a patient group.
Transfer to Patient Group	Transfer the list of patients to an existing patient group.
Download Patients	Download the list of patients.
Download Patient Details	Download patient details for the patients to an Excel spreadsheet or a Rich Text Format file. The application must be set up to allow you to download patient details.
Reports	View a report for the patients; you can run only one report definition at a time. Report definitions are listed only if they are compatible with the data configuration for the object that you are viewing.

Viewing a list of patients

After you drill down from a count of patients, the Patients page lists the patients in a table, with the patient ID in the first column. Other information in the table is determined by the data configuration.

The data on the Patients page is retrieved from the source data. Therefore, custom terms defined during the creation of analysis runs do not appear.

To view a list of patients from the Descriptive Analysis or Evaluative Analysis tabs:

1. Select the **Descriptive Analysis** or **Evaluative Analysis** tab.
2. Select the row menu () for a run, and select **View Results**.

   Results of the run appear. Several columns include patient counts with links.
3. Click a link in one of the columns with patient counts, and select **View Patients** from the drop-down list.

   Demographic information as determined by the configuration appears. The column containing the patient ID contains a link for viewing patient details.

To view a list of patients from any page with a patient count:

- Click a patient count, and select **View Patients**.

  The Patients dialog box appears. The following options are available as links above the table of patients.

Option	Description
Create Patient Group	Create a patient group containing patients in the list. If you create a patient group and view the group on the Patient Groups tab, you can comment on individual patients when reviewing patient details.
Transfer to Patient Group	Transfer the list of patients in the list to an existing patient group.
Download Patient Details	Download patient details for the patients to an Excel spreadsheet or a Rich Text Format file, depending on your preferences. The application must be set up to allow you to download patient details. **Note:** The **Download** option, which also appears on this page, downloads only the displayed information.
Reports	View a report for the patients.
Single-patient	For each patient in the list, view a graph of data across time in

Timelines	single-patient timelines.
Multi-patient Timelines	(Available only from the results of analysis runs.) View a graph of data across time in multi-patient timelines.

Creating a patient group

Counts of patients appear on various pages, including reports and the results of descriptive analysis and evaluative analysis runs. You can save a list of patients who are of interest to you by creating a patient group. You can use patient groups in analysis runs and reports.

For information about how to create a patient group from Browse Data, see Browsing patient data and creating a patient group.

1. Select the **Descriptive Analysis** or **Evaluative Analysis** tab.
2. Select the row menu () for an analysis run, and select **View Results**.

   Results of the run appear. Several columns include patient counts with links.
3. Click a link in one of the columns with patient counts, and select **Create Patient Group** from the drop-down list.

   The Create Patient Group dialog box appears.
4. Fill in the fields, and click **Create**.

Downloading data

On many pages, you can download the contents of the table to one of the following file formats:

- Comma-separated file (CSV).
- Tab-delimited file (TXT).
- Microsoft Excel spreadsheet (XLS).
- SAS Version 5 transport file (XPT).
- SAS data step definition (SAS).
- Portable document format (PDF).

Prerequisites

- Configure the Microsoft Internet Explorer browser for downloading data and printing graphs.
- Download third-party tools as needed, such as:
  - Adobe Acrobat Reader.
  - Microsoft Excel.
  - SAS System Viewer, for XPT files.
  - Base SAS, for SAS files.

To download data:

1. On a page with tabular data, click **Download**.

   Note: If the page has a **Select Rows** link, you can click Select Rows and download data for only the rows you select.

   The Download dialog box appears.

2. Fill in the fields. You can limit the number of rows to download.

3. Click **OK**.

Printing data

On many pages, you can print the data in a table.

Prerequisites

Configure the Microsoft Internet Explorer browser for downloading data and printing graphs.

To print data:

1. On a page with tabular data, click **Print**.

   Note: If the page has a **Select Rows** link, you can click Select Rows and print data for only the rows you select.

   The Print Table dialog box appears, displaying the data to be printed. Additionally, the Print dialog box appears.

2. In the Print dialog box, adjust printing settings as necessary, and click **Print**.

Selecting columns and rows

On many pages with tabular data, the **Columns** link appears. Click the link to open the Columns dialog box, where you can perform the following tasks:

- Add or remove columns from the table.
- Set the sort order of the columns.

Configuring the Microsoft Internet Explorer browser for downloading data and printing graphs

The following procedure is required if you plan to download data and print graphs and other outputs that use color. You can perform these steps at any time and need to do so only once.

1. Open the Microsoft Internet Explorer browser.
2. Select **Tools > Internet Options > Advanced** tab.
3. In the Printing section, select **Print background colors and images**. This setting allows the application to print the color key along with a graph, and to print the shading for table column headers and grid lines.
4. In the Security section, deselect **Do not save encrypted pages to disk** so that the application can download data.
5. Click **OK**.

Troubleshooting: Opening a file after downloading data

Due to a limitation in the Microsoft Internet Explorer browser, you might not be able to use the Open button in the File Download dialog box when you download data. If you encounter problems opening a file, perform the following steps.

- Save the file, and then open it.

  or

  Clear the Microsoft Internet Explorer cache by selecting **Tools > Internet Options > General** tab, and then click **Delete Files**.

Selecting multiple rows

1. On a page with tabular data, click **Select Rows**.
2. Select one or more rows to which to apply an action, and then click one of the following links:
    - Print
    - Download
    - Publish
    - Delete
3. To return to the previous page, click **Back**.