

# Atlas User Process Guide

Atlas is an open-source application that provides a unified interface to patient-level data and analytics.

## Table of Contents

Atlas User Process Guide .....	1
Atlas Navigation Features .....	2
Using Data Sources .....	2
Using Search.....	4
Using Concept Sets .....	5
Using Cohort Definitions .....	8
Using Characterizations .....	10
Using Cohort Pathways.....	11
Using Incidence Rates .....	12
Using Profiles .....	13
Using Estimation .....	14
Using Prediction.....	16
Using Jobs .....	18
Using Configuration.....	18
Using Feedback.....	18
Frequently Asked Questions .....	18
OHDSI Video Tutorials.....	18
Additional Resources .....	18
Glossary of Terms.....	19
Table of Figures .....	21



## Atlas Navigation Features

The left-hand navigation will direct your search for different analytic options (see Figure 1).

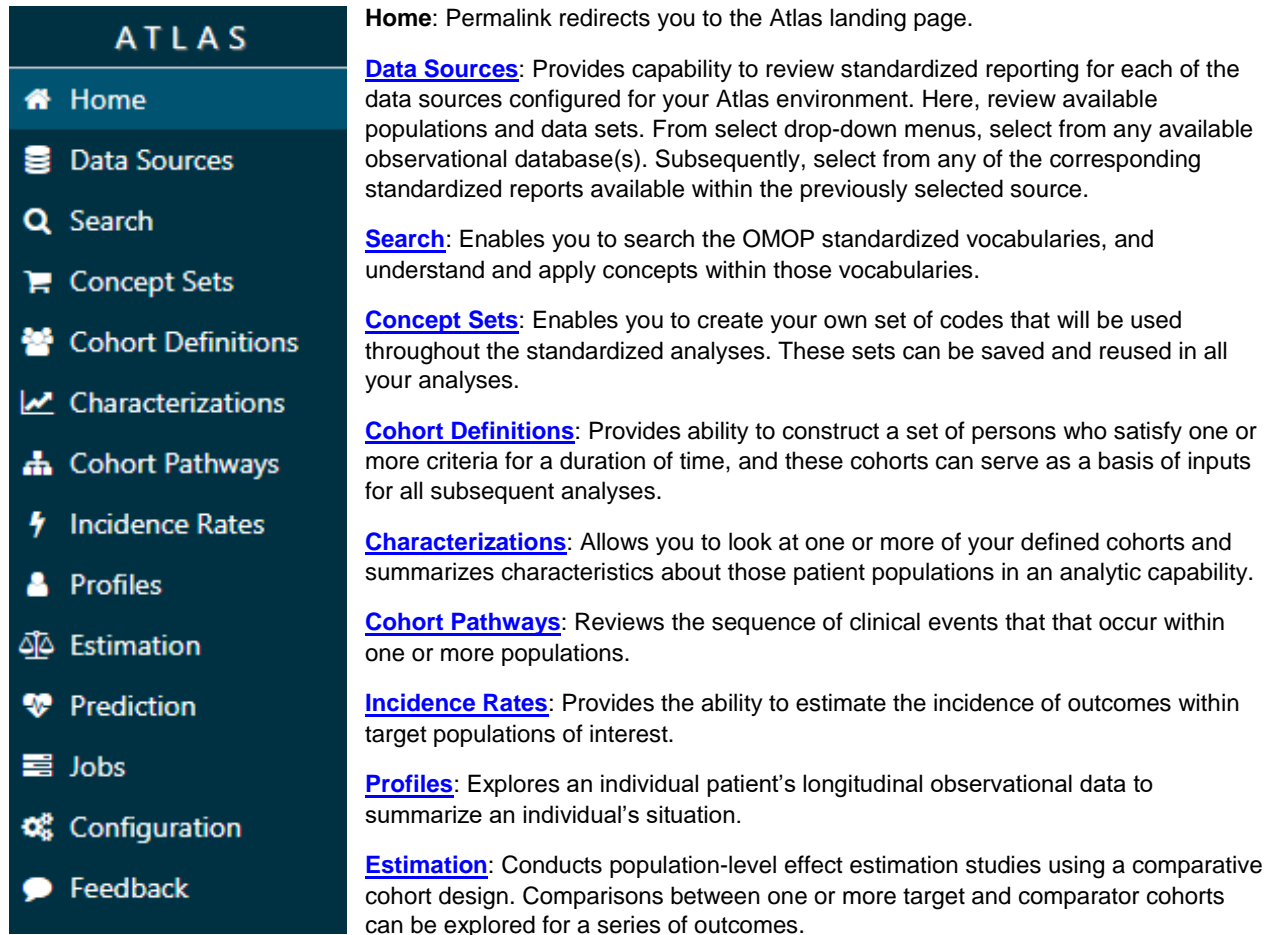


Figure 1: Atlas Navigation

given target exposures.

**Jobs:** Details queries and analyses. This list displays jobs that are running in the background for long-running jobs or reports.

**Configuration:** Displays data sources configured in the source configuration; while it is not possible to edit configurations here, it is possible to review options for the sources of the data available.

**Feedback:** Provides a method to contact Atlas administration regarding issues with the database, features, or enhancements.

## Using Data Sources

The Data Sources feature provides capability to review standardized reporting for each of the data sources configured for your Atlas environment. From here, review available populations and data sets.

Select **Data Sources** in the left-hand navigation.

1. From the first drop-down menu, select a **Source**. This has been configured for your Atlas environment.
2. From the second drop-down menu, select a **Report**. This displays a series of standardized reports.

Selecting the Report “**Dashboard**” will display:

- The source name and number of persons within the database
- Gender and age demographic charts
- Cumulative observation graph displaying the amount of time observed in the database
- Persons with Continuous Observation graph displaying the number of persons observed within a given month

Hover a mouse over a chart element in the dashboard to display additional quantitative information.

Several reports from the drop-down menu present a series of graphs. For every graph, hover a mouse over a table or chart element to render a pop-up that displays additional details and information. Some Report elements display Tree maps. Tree map representation shows a number of boxes, the size of which represents the prevalence of a condition concept within the database and the color of which represents its intensity (by number of records per person who have that concept).

1. Hover over a box within the tree map to display a record’s detailed information
2. Click on a box within the tree map to load a drilldown report with additional information about prevalence, type, and age at first occurrence

Next to the tab labeled “**Treemap**” is a tab for “**Table**.” Click the Table tab to represent the data in a tabular view (see Figure 2).

1. Click on a row of data to display additional information about the concept
2. Click the button “Show columns” to adjust which columns display
3. Click the button “Copy” to copy the concept information to your clipboard
4. Click the button “CSV” to export the information to an Excel spreadsheet

Concept Id	Name
0	No matching concept
9202	Outpatient Visit
9201	Inpatient Visit

Figure 2: Visit Report, Table Tab

Of the reports available under **Data Sources**:

Select the Report “**Data Density**” to display:

- A line report for every data domain supported
- A line report of records per person, enabling you to understand whether or not the data density seems to be growing as a function of the total data, or as a function of the amount of data per person
- A Concepts per Person line box plot, indicating distribution of distinct information per person across selected data domains

Select the Report “**Person**” to display a series of graphs, which include year of birth, gender, race, and ethnicity. These graphs enable an overview of demographic distribution and observe any potential data outliers.

Select the Report “**Visit**” to display a tree map representation showing the prevalence and density of information associated with different types of visits within the data source. To search for a specific visit type, click on the tab labeled “Table.” In the upper right-hand corner of the Table, type a query in the text field next to “Filter.”

Select the Report “**Condition Occurrence**” to display a tree map representation showing a number of boxes, the size of which represents the prevalence of a condition concept within the data source, and the color of which represents its intensity (by number of records per person who have that concept). This information is also displayed in the “Table” tab; click on any row in the Table to drill down to more information on a specific condition, specifically the concept prevalence over time, age information within the population, and frequency of diagnosis. To search for a specific condition, type a query in the text field next to “Filter” in the upper right-hand corner of the Table.

Select the Report “**Procedure**” to display a tree map representation showing the prevalence of all procedure concepts within the data source, the size of which represents the prevalence of a procedure concept within the data source, the color of which represents its intensity.

Select the Report “**Drug Exposure**” to display a tree map representation showing the prevalence and number of records per person the concept represents. Hover your mouse of a concept to display additional information (see Figure 3). As shown in Figure 3, the ingredient “Ibuprofen” within drug category “Musculo-Skeletal System Muscle Relaxants” indicates that the concept “ibuprofen 400 MG oral tablet” was prevalent in 0.85% of the population (98,790 individuals in this data source), and on average they have 16.03 records of ibuprofen 400 MG. Click on the box to open a drilldown report of prevalence, type, and age at first occurrence.

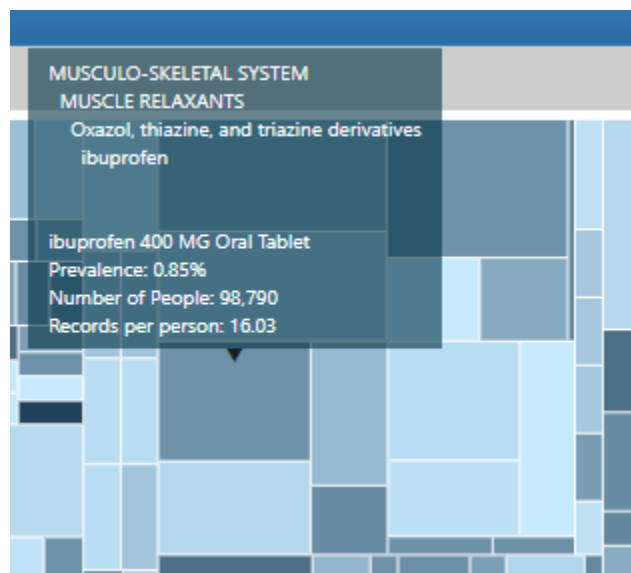


Figure 3: Tree map, Drug Exposure ibuprofen 400 MG

Select the Report “**Drug Era**” to display a tree map representation showing information about the ingredients used in continuous persistence exposure. Within this report, drugs have been aggregated to the ingredient level and represents a span of time that a person is exposed to a particular drug. Use the Tree map tab or the Table tab to review an overview of this information, or click on any concept to see a drilldown report with more information.

Select the Report “**Measurement**” to display a tree map representation showing vitals, labs, and other measurements of patients within the data source, the size of which represents the prevalence of a condition concept within the data source, and the color of which represents its intensity (by number of records per person who have that concept). This information is also displayed in the “Table” tab; click on any row in the Table to drill down to more information on a specific measurement, specifically the concept prevalence over time, age information within the population, and results of measurements. To search for a specific measurement type, type a query in the text field next to “Filter” in the upper right-hand corner of the Table.

(See [OHDSI Data Sources Video Tutorial](#))

## Using Search

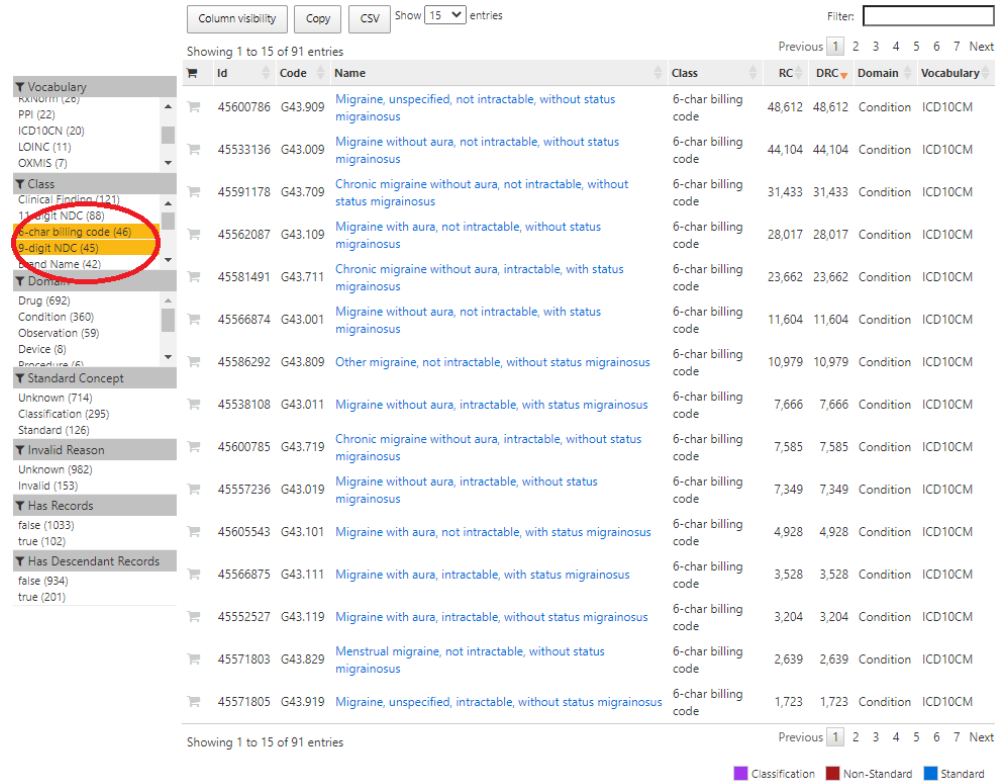
The Search tool enables you to search the OMOP standardized vocabularies and understand and apply concepts within those vocabularies.

Select **Search** from the left-hand navigation menu.

1. At the top, select the tab to “**Search**”

- In the **Search** field, enter words to find related concepts. For example, type “Migraine” in the search bar to render results that can be filtered. To filter, use the navigation to the left of the search results and click on items to add to the filtering criteria (see Figure 4).

The search feature can query conditions (e.g. migraine, hypertension, etc.), drugs (e.g. ACE inhibitor, Tylenol, etc.), vitals (e.g. height, weight, temperature, etc.), observations, procedures, devices, measurements, ICD-10 codes, etc.



Column visibility Copy CSV Show 15 entries Filter:

Showing 1 to 15 of 91 entries

	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
▼ Vocabulary								
▼ RXNorm (26)								
▼ PPI (22)								
▼ ICD10CN (20)								
▼ LOINC (11)								
▼ OXMIS (7)								
▼ Class								
▼ Clinical Finding (121)								
▼ 11-digit NDC (68)								
▼ 6-char billing code (46)								
▼ 8-digit NDC (45)								
▼ Brand Name (42)								
▼ Domain								
▼ Drug (692)								
▼ Condition (360)								
▼ Observation (59)								
▼ Device (8)								
▼ Procedure (6)								
▼ Standard Concept								
▼ Unknown (714)								
▼ Classification (295)								
▼ Standard (126)								
▼ Invalid Reason								
▼ Unknown (982)								
▼ Invalid (153)								
▼ Has Records								
▼ false (1033)								
▼ true (102)								
▼ Has Descendant Records								
▼ false (934)								
▼ true (201)								
45600786	G43.909	Migraine, unspecified, not intractable, without status migrainosus	6-char billing code	48,612	48,612	Condition	ICD10CM	
45533136	G43.009	Migraine without aura, not intractable, without status migrainosus	6-char billing code	44,104	44,104	Condition	ICD10CM	
45591178	G43.709	Chronic migraine without aura, not intractable, without status migrainosus	6-char billing code	31,433	31,433	Condition	ICD10CM	
45562087	G43.109	Migraine with aura, not intractable, without status migrainosus	6-char billing code	28,017	28,017	Condition	ICD10CM	
45581491	G43.711	Chronic migraine without aura, intractable, with status migrainosus	6-char billing code	23,662	23,662	Condition	ICD10CM	
45566874	G43.001	Migraine without aura, not intractable, with status migrainosus	6-char billing code	11,604	11,604	Condition	ICD10CM	
45586292	G43.809	Other migraine, not intractable, without status migrainosus	6-char billing code	10,979	10,979	Condition	ICD10CM	
45538108	G43.011	Migraine without aura, intractable, with status migrainosus	6-char billing code	7,666	7,666	Condition	ICD10CM	
45600785	G43.719	Chronic migraine without aura, intractable, without status migrainosus	6-char billing code	7,585	7,585	Condition	ICD10CM	
45557236	G43.019	Migraine without aura, intractable, without status migrainosus	6-char billing code	7,349	7,349	Condition	ICD10CM	
45605543	G43.101	Migraine with aura, not intractable, with status migrainosus	6-char billing code	4,928	4,928	Condition	ICD10CM	
45566875	G43.111	Migraine with aura, intractable, with status migrainosus	6-char billing code	3,528	3,528	Condition	ICD10CM	
45552527	G43.119	Migraine with aura, intractable, without status migrainosus	6-char billing code	3,204	3,204	Condition	ICD10CM	
45571803	G43.829	Menstrual migraine, not intractable, without status migrainosus	6-char billing code	2,639	2,639	Condition	ICD10CM	
45571805	G43.919	Migraine, unspecified, intractable, without status migrainosus	6-char billing code	1,723	1,723	Condition	ICD10CM	

Showing 1 to 15 of 91 entries

Previous 1 2 3 4 5 6 7 Next

Classification Non-Standard Standard

Figure 4: Search with Filters

Click on any line item to open a more detailed view within the vocabularies. Opening an item in this way presents a new set of tabs across the top of the page:

- Details** presents Vocabulary ID, Concept ID, Concept Code, and other property values connected to the record
- Related Concepts** provides other vocabulary for similar terms that may specify or broaden the search
- Hierarchies** indicates parents and children of the concept within the OMOP vocabulary
- Record Counts** displays the source of the records as well as the quantity

Clicking on a record to review these details will automatically navigate to the left-hand “**Concept Sets**” section.

(See [OHDSI Search Video Tutorial](#))

## Using Concept Sets

The Concept Sets tool enables you to create your own set of codes that will be used throughout the standardized analyses. These sets can be saved and reused in all your analyses. Concept sets are foundational to subsequent tools and analyses in Atlas.

Select **Concept Sets** in the left-hand navigation. Concept sets can contain any set of concepts across any of the domains within the OMOP standardized vocabulary. These can be customized so that different terms can be expressed in one item. Some entries have already been created, and new ones can be added.

Click on an existing Concept Set to provide a list of Concept Names and Domains contained within the set. Each item in the list has a corresponding check mark aligning under Exclude, Descendants, and Mapped. Clicking the box under the checked fields will add (or remove) the concept ID/names from the concept set (see Figure 5).

The screenshot shows the 'Concept Set #51' interface. At the top, there is a search bar containing 'COVID Dx' and buttons for 'Save', 'Close', 'Copy', 'Optimize', and 'Delete'. Below this is a navigation bar with 'Concept Set Expression' selected, and other options like 'Included Concepts' (17), 'Included Source Codes', 'Explore Evidence', 'Export', and 'Compare'. A 'Show 25 entries' dropdown and a search field are also present. The main content is a table with 3 entries:

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
704996	OMOP4912978	Patient meets COVID-19 laboratory diagnostic criteria	Observation	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
702953	U07.1	Emergency use of U07.1   COVID-19	Condition	Unknown	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
37311061	840539006	Disease caused by 2019-nCoV	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Figure 5: Concept Set Expressions (COVID Concept Set)

From the Concept Sets main page, ensure you are in the **“List”** tab. Click the blue button for **“New Concept Set.”** Enter a title for the concept set in the text box and click the green **“Save”** icon to begin (see Figure 6).

The screenshot shows the 'Concept Sets' main page. The 'List' tab is selected and circled in red. A blue button labeled 'New Concept Set' is also circled in red. Below the navigation bar, there is a table with columns: Id, Title, Created, Modified, and Author. The table shows several entries, including 'NAFLD/NASH', 'PB', 'HE', 'HD', 'HE', 'Pr', and 'Ar'. A 'New Concept Set' dialog box is open, showing a text input field with 'Migraine' entered, circled in red. To the right of the input field is a green 'Save' icon, also circled in red. The dialog box has a title bar 'New Concept Set' and a navigation bar with 'Concept Set Expression' selected, and other options like 'Included Concepts' (0), 'Included Source Codes', 'Explore Evidence', 'Export', and 'Compare'. At the bottom, there is a 'Show 25 entries' dropdown and a search field.

Figure 6: New Concept Set, Title and Beginning

In the left-hand navigation of the page, select the **“Search”** function.

- *Note:* using the smaller “Search” bar under the New Concept Set tabs will not render results, as this smaller search bar searches within the concept set. At the start of the creation, the concept set is empty, so there is nothing within the concept set to search.

1. In the search bar, enter concepts to identify for the new Concept Set
2. Use the left-hand search filters to add parameters to the search (for example: Standard Concept or Domain)
3. Click on a search result
4. In the top-menu tabs, navigate to **"Hierarchy"** to view the term's parents, concept, and children
5. Within the Hierarchy, click on any shopping cart icon to add it to the New Concept Set; the shopping cart will turn orange when it is added to the set (see Figure 7)

The hierarchy displays vocabulary you may want to include and perhaps exclude from the new concept set. Select concepts to exclude from the set by clicking the shopping cart icon; this adds the concept to the set, but functional adjustments will be made later.

**Vocabulary > Concept**

Migraine without aura

Details Related Concepts **Hierarchy** Record Counts

VIEW: Full Hierarchy Parents Current Children

Parents

Column visibility Copy CSV Show 15 entries

Showing 1 to 2 of 2 entries

Vocabulary	Id	Code	Name	Class	RC	DF
SNOMED (2)	372887	81308009	Disorder of brain	Clinical Finding	62,327	2,1
Clinical Finding (2)	318736	37796009	Migraine	Clinical Finding	104,691	33

Has Records true (2)

Has Descendant Records true (2)

Current Concept

Showing 1 to 1 of 1 entries

Id	Code	Name	Class	RC
378735	56097005	Migraine without aura	Clinical Finding	94,782

Children

Column visibility Copy CSV Show 15 entries

Showing 1 to 4 of 4 entries

Vocabulary	Id	Code	Name	Class	RC
SNOMED (4)	413616	423279000	Refractory migraine without aura	Clinical Finding	15,952
Clinical Finding (4)	765701	437931000124100	Chronic migraine without aura, non-refractory	Clinical Finding	0
Has Records false (3)					
true (1)	41534990	124001000119104	Common migraine with status migrainosus	Clinical Finding	0
Has Descendant Records false (3)	4314161	425007008	Migraine without aura, not refractory	Clinical Finding	0

Figure 7: Concept Hierarchy, Adding Terms to New Concept Set

At the top of the page, above the search results and above "Vocabulary > Concept" is a breadcrumb link to the "New Concept Set" (previously labeled by you, as detailed in Figure 6). Click this breadcrumb link to return to the Concept Set Expression tab.

1. The Concept Set Expression table provides columns to Exclude, Descendants, and Maps

2. Click the check box “**Exclude**” for concepts to exclude from the search
3. Click the check box for “**Descendants**” to add concepts’ children to the set
4. Click the green “Save” icon near the top of the page to save the concept set

Navigate to the left-hand menu item for “Concept Sets” to display the newly created concept set among the other pre-made sets.

([See OHDSI Concept Sets Video Tutorial](#))

## Using Cohort Definitions

The Cohort Definitions tool provides the ability to construct a set of persons who satisfy one or more criteria for a duration of time, and these cohorts can serve as a basis of inputs for all subsequent analyses.

Select **Cohort Definitions** in the left-hand navigation. Some definitions have already been created, and new ones can be added.

To create a new Cohort Definition, click the blue button labeled “**New Cohort Definition**” in the top right-hand corner.

1. In the top text box, enter a title for the New Cohort Definition; it can be specific (for example, “New users of ACE Inhibitors who have a prior diagnosis of hypertension”)
2. Click the green Save icon to the right of the text box
3. To begin defining components of the Cohort Definition, establish the cohort criteria (what must be observed so that someone enters the cohort, subpopulation, and how someone leaves the cohort criteria) by the following:
  - a. Under the “Cohort Entry Events” section, select “**Add Initial Event**” from the drop-down menu and click on an event to add (see Figure 8)
  - b. After selecting an event (for example, “Drug Exposure”) click the blue down-arrow next to the criteria item, and select “import concept” from the drop-down menu, which will provide a list of available [Concept Sets](#)
  - c. Continue to use different drop-down selections to specify date or event sequence
  - d. Click the blue “**Add Attribute**” button in order to specify demographic criteria (e.g. gender, first exposure, age, etc.) to add (see Figure 8)
  - e. Continue to add all relevant attribute criteria that pertain to the cohort
  - f. Once the event criteria are established, click the green button to “Restrict initial events”
4. In the “Inclusion Criteria” section, click the green button for “**New Inclusion Criteria**”
5. In the text field that appears, describe an inclusion criteria (for example “have a prior diagnosis of hypertension”)
6. Click the blue button to “**Add criteria to group**” in order to add inclusion criteria (e.g. criteria era, condition inclusion, device exposure, etc.)
  - *Note:* any criteria can be removed by clicking the red button labeled “**Delete Criteria**” to the right of the criteria box.
7. In the “Cohort Exit” section, select from the drop-down menu that the event will persist until a selected end
8. When all criteria are set, click the green Save icon at the top of the page



The screenshot shows the 'Cohort #36' interface. At the top, the title 'ACE Inhibitors users who have a prior diagnosis of hypertension' is highlighted with an orange box. To the right of the title are icons for Save, Close, Copy, Link, and Delete. Below the title are tabs for 'Definition', 'Concept Sets', 'Generation', 'Reporting', 'Export', and 'Messages'. A text box below the tabs says 'enter a cohort definition description here'. The main area is divided into three sections: 'Cohort Entry Events', 'Inclusion Criteria', and 'Cohort Exit'. The 'Cohort Entry Events' section has a dropdown menu for '+ Add Initial Event'. The 'Inclusion Criteria' section has a 'New inclusion criteria' button. The 'Cohort Exit' section has an 'Event Persistence' dropdown. A sidebar on the right lists various event types: 'Add Condition Era', 'Add Condition Occurrence', 'Add Death', 'Add Device Exposure', 'Add Dose Era', 'Add Drug Era', 'Add Drug Exposure', 'Add Measurement', and 'Add Observation'. The 'Add Drug Exposure' option is circled in red. Below the sidebar, there is a configuration area for the event, with a dropdown menu for 'Any Drug' and a text box for the event name. The 'Add attribute...' button is also circled in red.

Figure 8: Adding Cohort Entry Event to Cohort Set

After creating the criteria for a Concept Set:

1. At the top, select the tab for “**Concept Sets**” to review the concepts within your definition to ensure they represent your concepts of interest
2. Select the “**Generation**” tab to generate the Cohort Definition on a specific data source
3. Select the “**Export**” tab to review a full description of what the Cohort Definition represents
  - a. Near the top of the “Export” page, select buttons to view the information in Text View, Graphical View, JSON, or SQL
  - b. Select the blue “**Copy to Clipboard**” button to copy the information
4. Select the “**Messages**” tab to review warnings or memos regarding potential errors or incomplete aspects of the defined search criteria; this tab highlights yellow when a new warning generates

At the top of the Cohort Definitions page is the text box indicating the title of the definition. To the far right of the text box is a series of action buttons (see Figure 9).

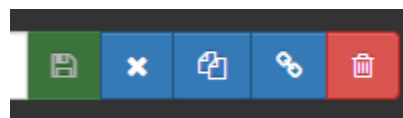


Figure 9: Cohort Definition Icons

The green Save icon saves changes to work. The blue X button closes the Cohort Definition and returns to the Cohort Definitions landing page. The blue double page icon creates a copy of the Cohort Definition. The blue link

button generates a pop-up with a direct link to share direct access to the Cohort Definition. The red trash can icon deletes the Cohort Definition.

(See [OHDSI Cohort Definitions Video Tutorial](#))

## Using Characterizations

The Characterizations tool allows you to look at one or more of your defined cohorts and summarizes characteristics about those patient populations in an analytic capability. This defines the process of generating a cohort-level descriptive summary statistics from person-level covariate data.

Select **Characterizations** in the left-hand navigation. Some characterizations have already been created, and new ones can be added.

On the Characterizations landing page, click on any of the listed existing Characterizations by name.

1. Select one or more Cohorts that will perform the characterization
2. In the Design tab, navigate to the Cohort Definitions section and click the blue button to **“Import” Cohort Definition** terms
3. From the pop-up, select items from the list of available [Cohort Definitions](#)
4. Navigate to the Feature analyses section and click the blue button to **“Import” features**
5. Use the left-hand filter navigation to filter results
6. Mark the check box next to all items to be included in the feature analysis (or click the green button to **“Select all”** (see Figure 10))
7. Click the blue **“Import”** button to add features

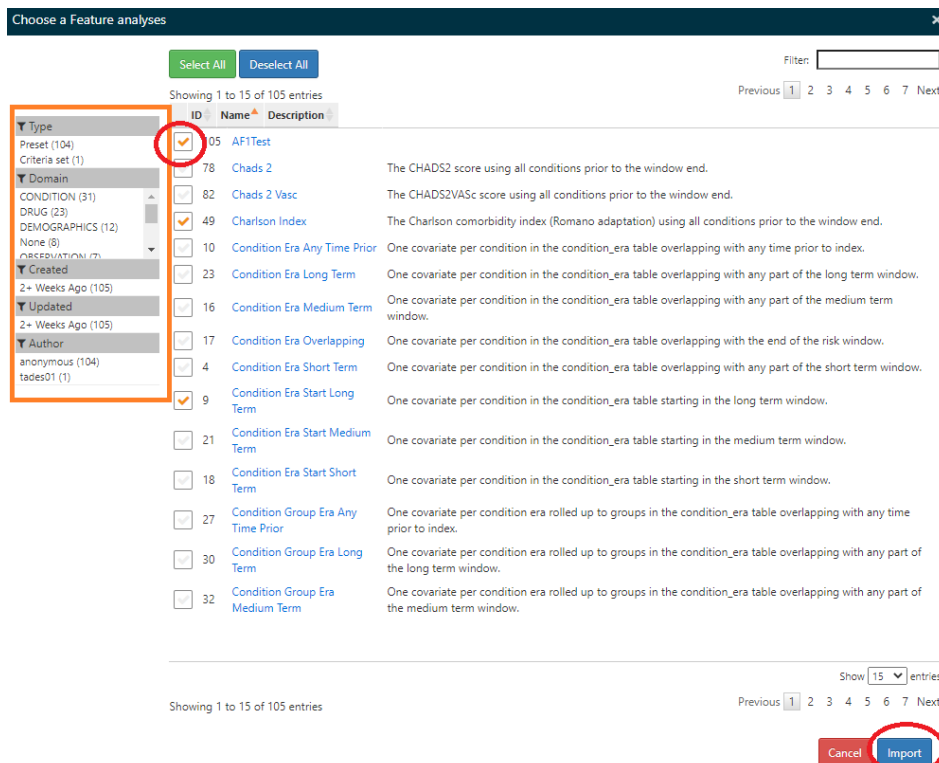


Figure 10: Adding Feature Analyses to Characterizations

Once characterizations have been established, run the execution:

1. At the top of the page, click the tab for “**Executions**”
2. Choose a source from which to run the data
3. Select the blue button to “**Generate**”
4. Once the generation is complete, click the adjacent blue button to “**View Latest Result**”

To select view options, navigate to the “Filter panel” section:

1. In the Cohorts drop-down menu, click the down-facing arrow to expand the menu
2. Each displayed cohort will have a black check mark to the right of the cohort name
3. Click on the check mark to remove or add the cohort name from the Executions display (see Figure 11)
4. Filter Analyses and Domains in the same manner

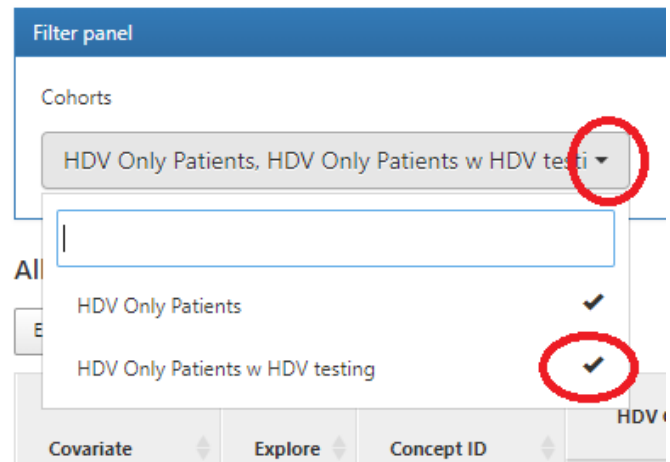


Figure 11: Filtering Cohorts in Characterizations > Executions

The Executions tab enables you to review All Prevalence Covariates, Demographics, and charts. Any section can be reviewed as an Excel spreadsheet by clicking the gray “**Export**” button under the corresponding section header.

([See OHDSI Characterizations Video Tutorial](#))

## Using Cohort Pathways

The Cohort Pathways tool reviews the sequence of clinical events that occur within one or more populations; it enables you to take one or more target populations and look at the sequence of events that occur based on specifications. By its process, it generates an aggregated sequence of transitions between a series of event cohorts among patients in a target cohort.

Select **Cohort Pathways** in the left-hand navigation. Some cohort pathways analyses may already exist, and new ones can be designed.

On the Cohort Pathways landing page, click on any of the listed existing pathways by name:

1. Click on one analysis, which will bring up a new page detailing the design specification
2. In the first “Target Cohorts” section, click the blue “**Import**” button to specify patient population(s) from the list of available [Cohort Definitions](#)
3. Click on a Cohort Definition’s name to add it to the Target Cohort section
4. In the second “Event Cohorts” section, click the blue “**Import**” button to specify the temporal sequence of events from the list of available Cohort Definitions
5. Click on a Cohort Definition’s name to add it to the Event Cohort section
6. Under Target and Event Cohorts, enter as many cohort selections as are relevant to the analysis
7. In the third “Analysis Settings” section, specify parameters to perform the analysis by selecting a number from each of the drop-down menus for the sequences you will analyze

After specifying Target Cohorts, Event Cohorts, and Analysis Settings:

1. Near the top of the page, select the gray “**Executions**” tab
2. In the Executions tab, select from a pre-set list of [Data Sources](#) configured for Atlas and click the corresponding blue “**Generate**” button
  - *Note:* a data generation may fail or time out; in such cases, the analysis parameters should be revised
3. Once the analysis is generated, click the adjacent blue “**View Latest Results**” button
4. The analysis results display on the report page that contains a Target Cohort legend and a sunburst plot; hover the mouse over an item in the sunburst plot will detail more specifics for the category

- a. The first (inner) ring of the sunburst plot represents the first-line therapy choice for patients
  - b. The subsequent lines radiating outward from the sunburst plot indicate second- or third- and combination sequence therapies within the population
5. Click any item in the sunburst plot to view additional path details, which then display in the right-hand panel

Near the top of the page, the third tab labeled “**Utilities**” provides an exportable code. The code enables you to create the design specification and share it with anyone else in the OHDSI community using Atlas. It enables a separate user to run the same analysis in their own environment.

([See OHDSI Cohort Pathways Video Tutorial](#))

## Using Incidence Rates

The Incidence Rates tool provides the ability to estimate the incidence of outcomes within target populations of interest.

Select **Incidence Rates** in the left-hand navigation. Some incidence rate analysis specifications may already be listed, and new analyses can be created.

On the Incidence Rates landing page, click on any of the listed existing analyses by name.

1. Click on one analysis, which will bring up a new page detailing the design specification for the incidence rate analysis
2. Review the listed groups under the category “**Target Cohorts**,” which represent the population at risk of experiencing some outcome
3. Review the listed groups under the category “**Outcome Cohorts**,” which represent the specific outcomes of interest to study within the target population

To create a new analysis, click the blue button labeled “**New Analysis**” in the top right-hand corner.

1. To add any new population groups to Target Cohorts, click the green “**Add Target Cohort**” button; click the red “**Add Outcome Cohort**” button to add new population groups to Outcome Cohorts (see Figure 11)
2. By clicking “Add Target Cohort” or “Add Outcome Cohort,” a pop-up containing all [Cohort Definitions](#) will appear; click on any definition name to add it to the list, and click the top-right x icon within the pop-up to close the pop-up menu when finished
3. To remove any population group from Target or Outcome Cohorts, click the black x icon to the left of the group name (see Figure 12)

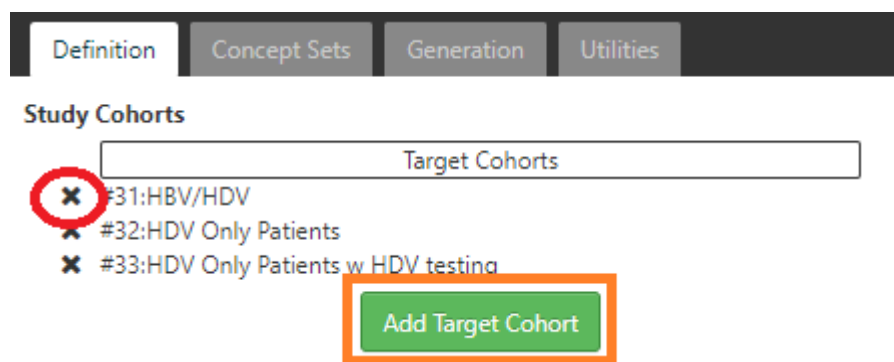


Figure 12: Add and Remove Target Cohorts

4. In the “Time At Risk” section, define a time window relative to the target cohorts’ start or end to determine if an outcome has occurred
  - a. From the drop-down menu, define a time at risk starts relative to either the start or end date of the target cohort (for example, in the case of “start date” plus “1” day, analysis reflects the day after you initiate HDV testing on HDV Only Patients)

- b. From the second drop-down menu, define a time at risk ends relative to either the start or end date of the target cohort (to continue from the previous example, analysis reflects the day after you initiate testing through to the end of testing)
5. To add a “Study Window,” click on the green “**Add Study Window**” box to enable a calendar date range filter to the analysis; this feature is optional
6. To add a subgroup stratification, click on the green “**New stratify criteria**” box to enable a text-entry label for specific groupings of the population (for example, “Gender=male” or “Age > 65”); this feature is also optional

After clearly defining incidence rates criteria for the new analysis, navigate to the top of the page and select the “**Generation**” tab:

1. Select which source(s) within Atlas to execute the analysis
2. Click the blue button labeled “**Generate**” to perform the analysis

The Generation page provides the results from selected sources. A general overview enables you to interpret:

- The source name (from within Atlas)
- The number of persons in the patient population
- The number of observed cases within the patient population
- A calculated incidence proportion based on the number of cases within the population
- A quantified time at risk (based in person years) with a corresponding incidence rate (in person years)

To see more detail from this data summary, click the blue “**Reports**” button to the right-hand side of the row. Reports displays:

- Summary statistics
- Impact of stratification criteria (refer to step 6, above)
- A tree map representation highlighting impact of stratification criteria together
  - Box size correlates to proportion of the population
  - Box color represents incidence rate gradient

At any time, select the green Save icon to save search or the blue x icon to close the search and return to the Incidence Rates landing page (refer to Figure 9).

[\(See OHDSI Incidence Rates Video Tutorial\)](#)

## Using Profiles

The Profiles tool displays an individual patient’s longitudinal observational data to summarize an individual’s situation.

Select **Profiles** in the left-hand navigation.

1. From the drop-down menu, select the database(s) from which you’ll be looking at data
2. Enter the individual’s ID for that database in the text field and press “Enter” on your keyboard to run the search
3. The graphical display for the patient will detail:
  - Patient demographics (age, gender, etc.)
  - Number of events
  - Data over time (day zero (0) representing first observation through end of continuous observation)
  - Swimlane data indicating timestamped clinical events (conditions, drugs, devices, observations, procedures, events, etc.)
4. Hover your mouse over any gray dot in the graphical display to see more detail about the specific data element

There are many different ways to see more detail within the patient records:

1. Within a swimlane, click your mouse and drag to the left or right to specify a period of time, which will zoom in on data in the specified time period

2. In the left-hand filter navigation, select a domain of interest under Doman Facet
3. Use the “**Filter**” text field on the right-hand side of the page between the graph and the tabular data to search for specific concepts

Deselecting a domain or double-clicking within the swimlane display will return to the graph’s starting point showing all information for the patient.

To color-code data in the swimlane graph:

1. Select the blue paintbrush icon near the top of the page (see Figure 13)
2. In the pop-up menu, use the “**Filter**” text field on the right-hand side to search for a particular concept (e.g. “Lisinopril” for all occurrences of this particular drug)
3. Click on the concept and then select a colored paintbrush icon near the top of the pop-up (e.g. green)
4. As desired, search for additional conditions or concepts (e.g. hypertension) and select a corresponding color (e.g. purple)
5. To close the pop-up, click the x icon in the top right-hand corner of the pop-up
6. Review the swimlane visualization graph with color-coded instances (e.g. diagnoses of hypertension in purple coinciding with administration or prescriptions for Lisinopril in green, etc.)

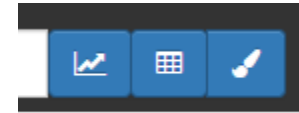


Figure 13: Patients Display Buttons

All data displayed visually in the graph is represented in tabular form beneath it. The tabular data contains:

- Concept ID
- Concept Name
- Domain
- Start and End date associated with each event

(See [OHDSI Profiles Video Tutorial](#))

## Using Estimation

Estimation enables you to conduct population-level effect estimation studies using a comparative cohort design. Comparisons between one or more target and comparator cohorts can be explored for a series of outcomes.

Select **Estimation** in the left-hand navigation. Some population-level effect estimations may already be listed, and new estimations can be created.

1. Click the blue “**New Population Level Effect Estimation**” button in the top right-hand corner
2. In the first section labeled “Comparative Cohort Settings,” click the blue “**+ Add Comparison**” button; this defines target and comparator cohorts (defining exposures, outcomes of interest, and negative control outcomes for empirical calibration)
  - a. In the first “Comparison” section, establish target and comparator cohorts
  - b. Click the blue folder icon to the right of each cohort field (see Figure 14)
  - c. In the pop-up menu, select a previously defined [Cohort Definition](#) from the list to add it

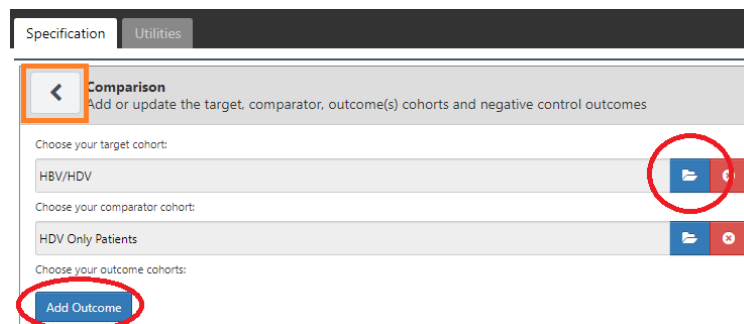


Figure 14: Adding Comparison Cohorts in Estimation Analyses

- d. Click the blue **“Add Outcome”** button (see Figure 14) and select a Cohort Definition from the pop-up menu to add it to the Outcomes list
  - e. In the field for negative control outcomes, click the blue folder icon to the right of the field and select a negative control from the pop-up of previously defined [Concept Sets](#)
  - f. In the Covariate selection section, include or exclude covariate concepts by clicking on the corresponding blue folder icon and selecting a concept from the pop-up of previously defined Concept Sets
  - g. Click the gray **“<”** button near the top of the page (see Figure 14) to return to the Estimation landing page and continue defining cohorts for analysis
3. In the second section labeled **“Effect Estimation Analysis Settings,”** click the blue **“+ Add Analysis Settings”** button; this establishes the ways statistical models will be applied to the specified comparison in **“Comparative Cohort Settings”**
    - a. In the first **“Study Population”** section, indicate parameters for study by entering dates in the text field (first two variable options) or selecting from the drop-down menus (remaining variable options)
    - b. In the second **“Covariate Settings”** section, a default recommended set of covariates has been pre-selected; view or edit these settings by following the **“Click to view details”** text link at the top of the section (see Figure 15)

**Covariate Settings**

Using OHDSI covariates for propensity score model. [\(Click to view details\)](#)

	Gender	Age	Age Groups	Race	Ethnicity	Index Year	Index Month	Prior Observation Time	Post Observation Time	Time In Cohort	Index Year & Month
Demographics	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Time bound covariates**

Set the time windows for the time bound covariates in days relative to the cohort index

	Any Time Prior	Long Term	Medium Term	Short Term	End Days
Time Windows	All Time	-365	-180	-30	0

**Set the time bound era covariates**

Domain	Any Time Prior	Long Term (-365 days)	Medium Term (-180 days)	Short Term (-30 days)	Overlapping	Era Start		
						Long Term (-365 days)	Medium Term (-180 days)	Short Term (-30 days)
Condition	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Condition Group	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Drug	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Drug Group	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Set the time bound covariates**

Domain	Any Time Prior	Long Term (-365 days)	Medium Term (-180 days)	Short Term (-30 days)	Distinct Count		
					Long Term (-365 days)	Medium Term (-180 days)	Short Term (-30 days)
Condition	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Condition - Primary Inpatient	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 15: Covariate Selection in Estimation

- c. In the third **“Time At Risk”** section, specify at-risk window parameters from the drop-down menus to create a time frame; this would convey, for example, events occurring sometime after the start of treatment up to the point a patient was no longer continuously exposed
- d. In the fourth **“Propensity Score Adjustment”** section, adjust for the predicted probability of assigned treatment which can trim patients on the basis of extreme values

1. Within the section are subsections; click the blue “Additional Covariate Settings” header to expand the menu and input specific covariate IDs to include or exclude in the Propensity Score Adjustment
2. Click the blue “Control Settings” header to adjust additional control settings
3. Click the blue “Prior” header to adjust for LASSO logistic
4. Clicking the blue header for each of these categories will expand and collapse the menu
- e. In the fifth “Outcome Model Settings” section, specify the statistical model:
  - Logistic regression: the odds of an outcome occurrence
  - Poisson regression: the rate of events
  - Cox Proportional Hazards: hazards of time to onset of an event
- f. Click the gray “<” button near the top of the page (compare to Figure 14) to return to the Estimation landing page and continue defining cohorts for analysis
4. In the third section labeled “Evaluation Settings,” establish how negative controls will try to eliminate systematic error in analysis and to empirically calibrate final summary statistics.

Once the analysis is specified, scroll to the top of the Estimation page and click the “Utilities” tab (refer to Figure 16 for comparable visualization)

1. In the presented table, review all specified analysis settings
2. In the text field below the table, type a name for the study package and click the blue “Download” button
3. The R-package downloads as a .zip file which can be uploaded into an Atlas studio
4. Click the “Export” button near the top of the page (refer to Figure 15 for comparable visualization) to review the JSON expression, which can be shared with anyone else in the OHDSI community using Atlas; it will enable a separate user to run the same analysis in their own environment

[\(See OHSDI Video Tutorial for Estimation\)](#)

## Using Prediction

The Prediction tool allows you to apply machine-learning algorithms to conduct patient-level prediction analyses. From here, it is possible to predict an outcome within any given target exposures

Select **Prediction** in the left-hand navigation. Some population-level effect estimations may already be listed, and new patient-level predictions can be created.

1. Click the blue “New Patient Level Prediction” button in the top right-hand corner of the screen
2. In the first “Prediction Problem Settings” segment, specify the target cohort representing the at-risk population you want to predict and specify the outcome cohort representing the outcome experienced in the target population that you are looking for
  - a. Within “Prediction Problem Settings,” click the blue “+ Add Target Cohort” button
  - b. Select a population (or multiple populations) from the pop-up of previously defined [Cohort Definitions](#)
3. In the second “Analysis Settings” segment, specify one or more machine learning algorithms to apply to the study population
  - a. Click the blue “+ Add Model Settings” to specify the algorithm type(s) to be applied
  - b. Click the blue “+ Add Covariate Settings” to specify covariates used as predictors in that machine learning algorithm
    - *Note:* in the pop-up, specify additional features of interest by checking or unchecking boxes
  - c. Click the blue “+ Add Population Settings” to represent the time at-risk and other considerations for defining the population
    - *Note:* this represents a span of time relative to the target cohort where you are looking for outcomes
    - *For example:* you are searching for patients that are new users of a drug and are looking for outcomes that occur within one year following the start of treatments
    - *Note:* additionally, specify the population criteria if, for example, patients without time-at-risk should be removed from the analysis, etc.



4. In the third “Execution Settings” segment and the fourth “Training Settings” segment, define your settings to execute the analysis
  - *Note:* the model will be trained on some training data set and a holdout independent test set will be used to provide internal validation of the models

Once the analysis is specified, scroll to the top of the Prediction page and click the “**Utilities**” tab (see Figure 16). If there is an error in the formula or analysis, the Utilities tab will describe the nature of the error before presenting results; in cases like this, click the “Specification” tab to return to the analysis parameters and make adjustments.

The screenshot shows the Prediction Results page with the Utilities tab selected. The page includes a navigation bar with 'Specification' and 'Utilities' tabs, and buttons for 'Review & Download', 'Import', and 'Export'. The main content area is titled 'Review & Download' and contains a 'Review Full Study Specification' section with a table of analysis settings. Below the table is a 'Download Study Package' section with a text input field for the package name and a 'Download' button.

Target Cohort	Outcome Cohort	Model Name	Model Settings	Covariate Settings
AF Condition (1)	Stroke (1)	LassoLogisticRegressionSettings	{"variance": "5", "seed": "null"}	"attr_class": "covariateSettin...

Showing 1 to 1 of 1 entries

Package name...

Please **Note:** The package name should consist of alphanumeric characters only. No spaces or special characters are permitted.

Download

Figure 16: Prediction Results

1. In the presented table, review all specified analysis settings
2. In the text field below the table, type a name for the study package and click the blue “**Download**” button (see Figure 16)
3. The R-package downloads as a .zip file which can be uploaded into an Atlas studio
4. Click the “Export” button near the top of the page to review the JSON expression (see Figure 16, marked in orange), which can be shared with anyone else in the OHDSI community using Atlas; it enables a separate user to run the same analysis in their own environment

(See [OHDSI Video Tutorial for Prediction](#))

## Using Jobs

The Jobs tool details queries and analyses. This list displays jobs that are running in the background for long-running jobs or reports.

Select **Jobs** in the left-hand navigation. The table displays all jobs run, the status of the job, who ran the job, and when.

Click the green “**Refresh Jobs**” button in the top right-hand corner of the Jobs page. This refreshes the content.

Click the gray “**CSV**” button at the top of the Jobs page near the left-hand side. This downloads an Excel spreadsheet of all jobs and the corresponding data displayed in the table.

## Using Configuration

The Configuration tool displays data sources configured in the source configuration; while it is not possible to edit configurations here, it is possible to review options for the sources of the data available.

Select **Configuration** in the left-hand navigation. The list displays all data sources configured for your Atlas.

*Note:* it is not possible to edit the source configuration; edits must be done directly in the database.

## Using Feedback

The Feedback feature provides a method for you to contact Atlas administration regarding issues with the database, features, or enhancements. Select **Feedback** in the left-hand navigation and follow instructions in order to submit issues.

## Frequently Asked Questions

### Q: Why are some features broken/not working?

Try switching web browsers. Google Chrome is the recommended browser for Atlas.

## OHDSI Video Tutorials

OHSDI has provided extensive video tutorials demonstrating features and functions of Atlas.

- [Using Data Sources: Dashboard](#)
- [Using Search](#)
- [Using Concept Sets](#)
- [Using Cohort Definitions](#)
- [Using Characterizations](#)
- [Using Cohort Pathways](#)
- [Using Incidence Rates](#)
- [Using Profiles](#)
- [Using Estimation](#)
- [Using Prediction](#)

## Additional Resources

- [OHDSI's Atlas Wiki](https://github.com/OHDSI/Atlas/wiki) (permalink: <https://github.com/OHDSI/Atlas/wiki>)
- [Book of OHDSI](https://ohdsi.github.io/TheBookOfOhdsi/) (permalink: <https://ohdsi.github.io/TheBookOfOhdsi/>)
- [Athena Domain Wiki](https://athena.ohdsi.org/search-terms/start) (permalink: <https://athena.ohdsi.org/search-terms/start>)

## Glossary of Terms

### **Bias**

The expected value of the error (the difference between the true value and the estimated value).

### **Causal effect**

What population-level estimation concerns itself with. One definition equates a “causal effect” as the average of the “unit-level causal effects” in a target population. The unit-level causal effect is the contrast between the outcome had an individual been exposed and the outcome had that individual not been exposed (or been exposed to A as against B).

### **Characterization**

Descriptive study of a cohort or entire database.

### **Cohort**

A set of persons who satisfy one or more inclusion criteria for a duration of time.

### **Concept**

A term (with a code) defined in a medical terminology (e.g., SNOMED CT).

### **Concept set**

A concept set is an expression representing a list of concepts that can be used as a reusable component in various analyses.

### **Common Data Model (CDM)**

A convention for representing healthcare data that allows portability of analysis (the same analysis unmodified can be executed on multiple datasets).

### **Condition**

A diagnosis, a sign, or a symptom, which is either observed by a provider or reported by the patient.

### **Covariate**

Data element (e.g., weight) that is used in a statistical model as independent variable.

### **Device**

A foreign physical object or instrument which is used for diagnostic or therapeutic purposes through a mechanism beyond chemical action. Devices include implantable objects (e.g. pacemakers, stents, artificial joints), medical equipment and supplies (e.g. bandages, crutches, syringes), other instruments used in medical procedures (e.g. sutures, defibrillators) and material used in clinical care (e.g. adhesives, body material, dental material, surgical material).

### **Drug**

A Drug is a biochemical substance formulated in such a way that when administered to a Person it will exert a certain physiological effect. Drugs include prescription and over-the-counter medicines, vaccines, and large-molecule biologic therapies. Radiological devices ingested or applied locally do not count as Drugs.

### **Domain**

A Domain defines the set of allowable Concepts for the standardized fields in the CDM tables. For example, the “Condition” Domain contains Concepts that describe a condition of a patient, and these Concepts can only be stored in the condition\_concept\_id field of the CONDITION\_OCCURRENCE and CONDITION\_ERA tables.

### **Matching**

Many population-level effect estimation approaches attempt to identify the causal effects of exposures by comparing outcomes in exposed patients to those same outcomes in unexposed patients (or exposed to A versus B). Since these two patient groups might differ in ways other than exposure, “matching”

attempts to create exposed and unexposed patient groups that are as similar as possible at least with respect to measured patient characteristics.

**Measurement**

A structured value (numerical or categorical) obtained through systematic and standardized examination or testing of a person or person's sample.

**Measurement error**

Occurs when a recorded measurement (e.g., blood pressure, patient age, duration of treatment) differs from the corresponding true measurement.

**Negative control**

An exposure-outcome pair where the exposure is believed to not cause or prevent the outcome. Can be used to assess whether effect estimation methods produce results in line with the truth.

**Observation**

A clinical fact about a Person obtained in the context of examination, questioning or a procedure.

**Observation period**

The span of time for which a person is at-risk to have clinical events recorded within the source systems, even if no events in fact are recorded (healthy patient with no healthcare interactions).

**Outcome**

An observation that provides a focal point for an analysis. For example, a patient-level predictive model might predict the outcome "stroke." Or a population-level estimation might estimate the causal effect of a drug on the outcome "headache."

**Patient-level prediction**

Development and application of predictive models to produce patient-specific probabilities for experiencing some future outcome based on baseline characteristics.

**Population-level estimation**

A study into causal effects. Estimates an average (population-level) effect size.

**Positive control**

An exposure-outcome pair where the exposure is believed to cause or prevent the outcome. Can be used to assess whether effect estimation methods produce results in line with the truth.

**Procedure**

Activity or process ordered by, or carried out by, a healthcare provider on the patient to have a diagnostic or therapeutic purpose.

**Propensity score (PS)**

A single metric used in population-level estimation to balance populations in order to mimic randomization between two treatment groups in an observational study. The PS represents the probability of a patient receiving a treatment of interest as a function of a set of observed baseline covariates. It is most often calculated using a logistic regression model where the binary outcome is set to one for the group receiving the target treatment of interest and to zero for the comparator treatment.

**Selection bias**

A bias that occurs when the set of patients in your data deviates from the patients in the population in ways that distort statistical analyses.

**Self-controlled designs**

Study designs that compare outcomes during different exposures within the same patient.

**Sensitivity analysis**

A variant of the main analysis used in a study to assess the impact of an analysis choice over which uncertainty exists.

**Standard Concept**

A concept that is designated as valid concept and allowed to appear in the CDM.

**Visit**

The span of time a person continuously receives medical services from one or more providers at a care site in a given setting within the health care system.

**Vocabulary**

A list of words and often phrases, usually arranged alphabetically and defined or translated.

All definitions derived from [the Book of OHDSI](#) (open source). See Standardized Vocabularies from OHDSI [here](#).

## Table of Figures

Figure 1: Atlas Navigation.....	2
Figure 2: Visit Report, Table Tab .....	3
Figure 3: Tree map, Drug Exposure ibuprofen 400 MG .....	4
Figure 4: Search with Filters .....	5
Figure 5: Concept Set Expressions (COVID Concept Set) .....	6
Figure 6: New Concept Set, Title and Beginning .....	6
Figure 7: Concept Hierarchy, Adding Terms to New Concept Set .....	7
Figure 8: Adding Cohort Entry Event to Cohort Set .....	9
Figure 9: Cohort Definition Icons .....	9
Figure 10: Adding Feature Analyses to Characterizations .....	10
Figure 11: Filtering Cohorts in Characterizations > Executions .....	11
Figure 12: Add and Remove Target Cohorts .....	12
Figure 13: Patients Display Buttons.....	14
Figure 14: Adding Comparison Cohorts in Estimation Analyses.....	14
Figure 15: Covariate Selection in Estimation .....	15
Figure 16: Prediction Results.....	17