

PEDSnet Data Science Training



June 24, 2019

Overview

- Web sessions

Data Model & Vocabularies	April 2019
Data Request Fulfillment	May 2019
Web Application	June 2019
PEDSnet Standard R Framework	July 2019

- Hands-on workshop – August 01-02, 2019
- Use of network data and scientific projects

Atlas

Cohort Discovery



Atlas as a Self-Service Tool

- Data exploration – replaces Achilles
 - Cohort construction
 - Defined analyses
-
- Features are generally for people who understand data model and vocabularies
 - Best used for feasibility assessments for studies
 - We will focus today on cohort definition

Atlas Web Application

- Created by OHDSI
- Open source software – constantly in development
- Collaborative and all code available on Github
- Multi-institution effort
- Assumes data model in OMOP format

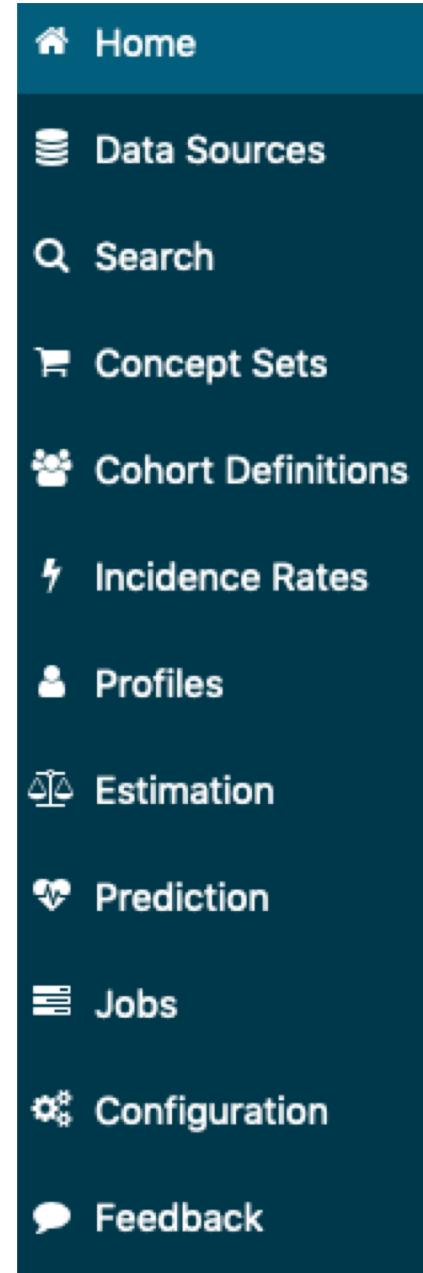


Atlas Components

- WebAPI
 - Java-based application designed to provide web services for interacting with one or more databases converted to OMOP CDM
 - Stores cohort definitions, concept sets, analysis specifications, etc
- Achilles
 - Tool that characterizes data
 - Index generation for better performance with Atlas Data Sources
- NodeJS
 - Allows use of all JavaScript dependencies

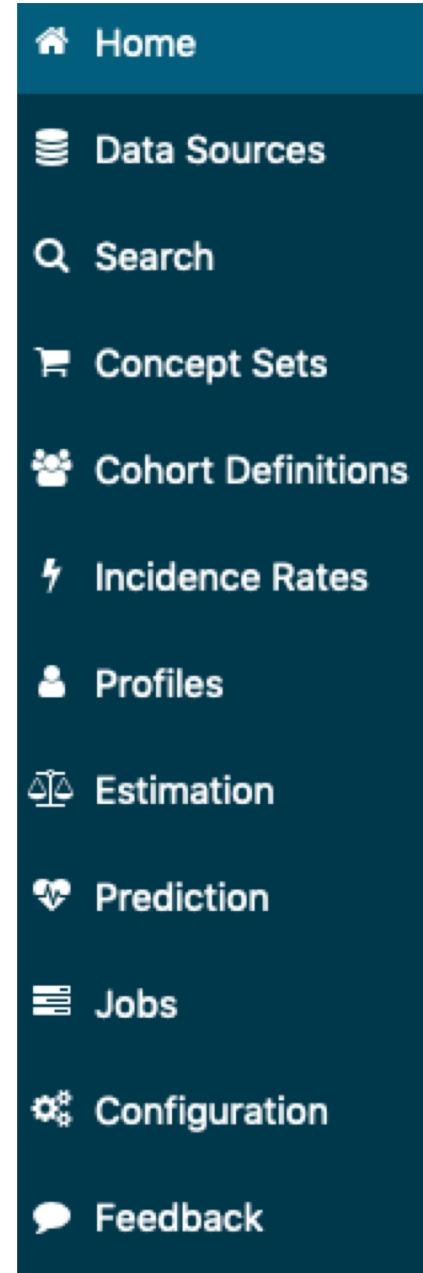
Atlas Features

- Data model navigation
- Vocabulary search and codeset creation
- Cohort construction and overviews
- Defined analyses using constructed cohorts



Atlas Features - Cohorts

- Codeset creation
- Definition based on criteria around an index date
 - Allows temporal associations
 - Multiple criteria across different domains
- Generate attrition table by inclusion criteria
- Visualizations available
- Saves cohorts that are created
- Import / Export cohorts with other networks
 - JSON, SQL

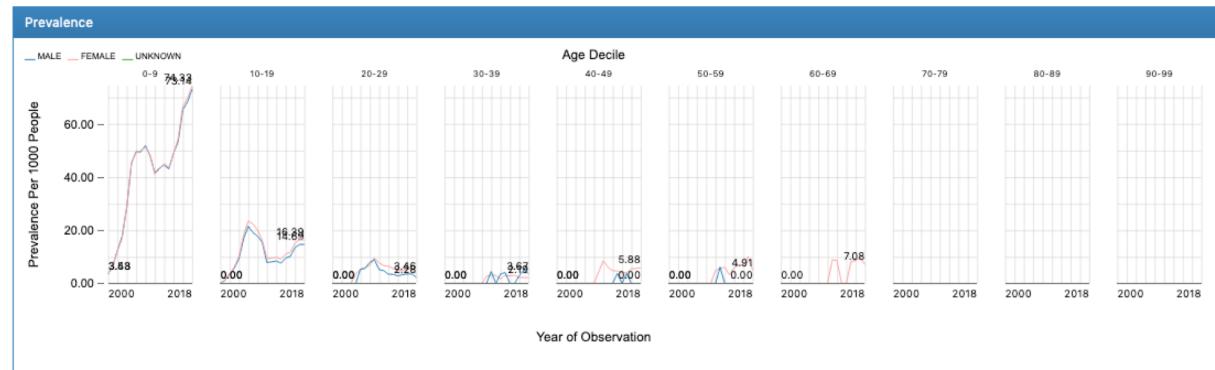


Data Source Exploration

Condition report(pedsnet_dcc)



Drilldown report(Acute upper respiratory infection)



Home

Data Sources

Search

Concept Sets

Cohort Definitions

Incidence Rates

Profiles

Estimation

Prediction

Jobs

Configuration

Feedback

Data Source Exploration

- Achilles-style descriptives and checks
- Provides overview of domain
(e.g. what is the overall diagnosis mix in PEDSnet?)
- Useful for single-value assessment
(e.g. how many patients received drug X?)
- Includes very abbreviated data quality tests (Achilles Heel)

Achilles Heel report(pedsnet_dcc)

Achilles Heel Results	
<input type="button" value="Column visibility"/>	<input type="button" value="Copy"/>
<input type="button" value="CSV"/>	Show 15 entries
	Filter: <input type="text"/>
Showing 1 to 15 of 68 entries	
Previous	<input type="button" value="1"/> <input type="button" value="2"/> <input type="button" value="3"/> <input type="button" value="4"/> <input type="button" value="5"/> Next
Message Type	Message
NOTIFICATION	Unmapped data over percentage threshold in:Procedure
NOTIFICATION	Unmapped data over percentage threshold in:Measurement
NOTIFICATION	Unmapped data over percentage threshold in:DrugExposure
NOTIFICATION	Unmapped data over percentage threshold in:Condition

Vocabulary Exploration

Q Search

Search Import

pneumonia

Column visibility Copy CSV Show 15 entries Advanced Options Filter:

Showing 1 to 15 of 3,626 entries

	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
	4045314	122120003	Streptococcus pneumoniae antigen assay	Procedure	0	1,645,754	Measurement	SNOMED
	4039553	117883007	Streptococcus pneumoniae group A antigen assay	Procedure	0	1,645,754	Measurement	SNOMED
	529072	798232	Streptococcus pneumoniae serotype 9V capsular antigen diphtheria CRM197 protein conjugate vaccine	Ingredient	473,228	955,527	Drug	RxNorm
	528988	798222	Streptococcus pneumoniae serotype 18C capsular antigen diphtheria CRM197 protein conjugate vaccine	Ingredient	473,227	955,526	Drug	RxNorm
	528986	798220	Streptococcus pneumoniae serotype 14 capsular antigen diphtheria CRM197 protein conjugate vaccine	Ingredient	473,226	955,525	Drug	RxNorm
	529044	798228	Streptococcus pneumoniae serotype 4 capsular antigen diphtheria CRM197 protein conjugate vaccine	Ingredient	473,226	955,525	Drug	RxNorm

Vocabulary
LOINC (1274)
SNOMED (766)
RxNorm Extension (484)
MedDRA (236)
CPT (1,100)
Class
Lab Test (897)
Clinical Finding (375)
LOINC Hierarchy (302)
LLT (184)
Condition (1127)
Measurement (1262)
Drug (791)
Observation (367)
Classification (576)
Standard Concept
Standard (1746)
Non-Standard (1304)
Classification (576)
Invalid Reason
Valid (3012)
Invalid (614)
Has Records
false (3402)
true (224)
Has Descendant Records
false (3259)
true (367)

Home Data Sources Search

Concept Sets Cohort Definitions Incidence Rates Profiles Estimation Prediction Jobs Configuration Feedback

PEDSnet A Pediatric Learning Health System

Vocabulary Exploration - Narrowing

- Search for terms (words) or codes
- Use left column to narrow search to
 - Standard concepts – what's in the database
 - Domains of interest
 - Terms in use

T Vocabulary
LOINC (1274)
SNOMED (766)
RxNorm Extension (484)
MedDRA (236)
...
T Class
Lab Test (897)
Clinical Finding (375)
LOINC Hierarchy (302)
LLT (184)
...
T Domain
Measurement (1262)
Condition (1127)
Drug (791)
Observation (367)
...
T Standard Concept
Standard (1746)
Non-Standard (1304)
Classification (576)
T Invalid Reason
Valid (3012)
Invalid (614)
T Has Records
false (3402)
true (224)
T Has Descendant Records
false (3259)
true (367)

pneumonia



Advanced Options

[Column visibility](#) [Copy](#) [CSV](#) Show 15
Filter:

Showing 1 to 15 of 129 entries

[Previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) ... [9](#) [Next](#)

Vocabulary

LOINC (1274)

SNOMED (766)

RxNorm Extension (484)

MedDRA (236)

Class

Lab Test (897)

Clinical Finding (375)

LOINC Hierarchy (302)

LLT (184)

Domain

Measurement (1262)

Condition (1127)

Drug (791)

Observation (367)

Standard Concept

Standard (1746)

Non-Standard (1304)

Classification (576)

Invalid Reason

Valid (3012)

Invalid (614)

Has Records

false (3402)

true (224)

Has Descendant Records

false (3259)

true (367)

	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
	255848	233604007	Pneumonia	Clinical Finding	221,017	594,689	Condition	SNOMED
	443410	312342009	Infective pneumonia	Clinical Finding	119,600	298,355	Condition	SNOMED
	257315	53084003	Bacterial pneumonia	Clinical Finding	61,578	112,481	Condition	SNOMED
	261326	75570004	Viral pneumonia	Clinical Finding	47,794	61,655	Condition	SNOMED
	4133224	278516003	Lobar pneumonia	Clinical Finding	8,331	31,712	Condition	SNOMED
	258180	430395005	Pneumonia due to Gram negative bacteria	Clinical Finding	3,004	31,289	Condition	SNOMED
	4273378	64667001	Interstitial pneumonia	Clinical Finding	258	18,101	Condition	SNOMED
	4231983	406595002	Infection due to Mycoplasma pneumoniae	Clinical Finding	160	16,128	Condition	SNOMED
	440431	46970008	Mycoplasma pneumonia	Clinical Finding	15,958	15,958	Condition	SNOMED
	4116487	300999006	Basal pneumonia	Clinical Finding	1,618	15,340	Condition	SNOMED
	4050869	233606009	Atypical pneumonia	Clinical Finding	10,252	10,252	Condition	SNOMED
	261324	34020007	Pneumonia due to Streptococcus	Clinical Finding	2,144	9,709	Condition	SNOMED

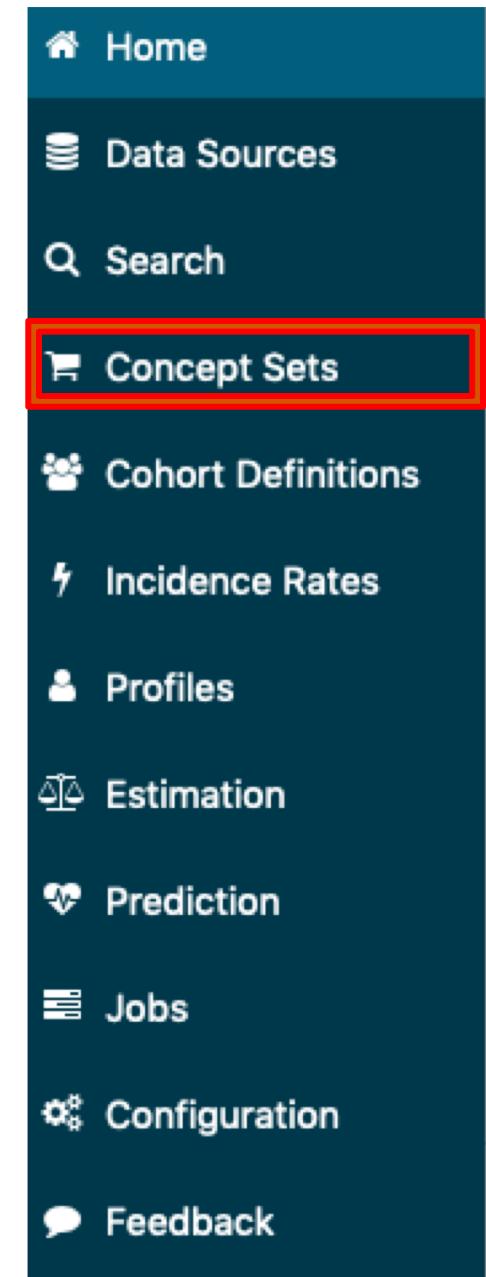
Vocabulary Exploration - Results

- Specific Terms
- RC = records using *this term exactly*
- DRC = records using *this term and its descendants*
- Note that these are record counts, not person counts
- Note the difference between has records and has descendant records

Name	Class	RC	DRC
Amoxicillin	Ingredient	3,373,941	9,092,982
Amoxicillin Oral Suspension	Clinical Drug Form	0	3,747,357
Amoxicillin 80 MG/ML	Clinical Drug Comp	0	3,469,630
Amoxicillin 80 MG/ML Oral Suspension	Clinical Drug	2,768,756	3,180,037
Amoxicillin / Clavulanate Oral Suspension	Clinical Drug Form	0	1,036,283
Amoxicillin 120 MG/ML	Clinical Drug Comp	0	691,971
Amoxicillin 120 MG/ML / Clavulanate 8.58 MG/ML Oral Suspension	Clinical Drug	592,627	691,971
Amoxicillin 50 MG/ML	Clinical Drug Comp	0	507,862
Amoxicillin 50 MG/ML Oral Suspension	Clinical Drug	461,657	482,386

Concept Sets

- Atlas name for codeset
- May be created, generated from list, or imported
- Fundamentally consists of a list of **concept_ids**
- Forms the basis for Atlas' interaction with the data
- May exist globally (prefix with initials) or just for a specific cohort



Concept Sets

- Start in Concept Set screen globally or in cohort
- Use Vocabulary Search to find terms of interest
 - Add by clicking on shopping cart
- Return to Concept Set to decide how term is used

This screenshot shows a table of search results for the term "pneumonia". The table has columns for Id, Code, Name, Class, RC, DRC, Domain, and Vocabulary. The results are filtered by SNOMED. The first few rows include:

Id	Code	Name	Class	RC	DRC	Domain	Vocabulary
255848	233604007	Pneumonia	Clinical Finding	221,017	594,689	Condition	SNOMED
443410	312342009	Infective pneumonia	Clinical Finding	119,600	298,355	Condition	SNOMED
257315	53084003	Bacterial pneumonia	Clinical Finding	61,578	112,481	Condition	SNOMED
261326	75570004	Viral pneumonia	Clinical Finding	47,794	61,655	Condition	SNOMED
4133224	278516003	Lobar pneumonia	Clinical Finding	8,331	31,712	Condition	SNOMED
258180	430395005	Pneumonia due to Gram negative bacteria	Clinical Finding	3,004	31,289	Condition	SNOMED

This screenshot shows a table of included concepts for the Renal Calculus Diagnosis concept set. The table has columns for Concept Id, Concept Code, Concept Name, Domain, Standard Concept Caption, and checkboxes for Exclude, Descendants, and Mapped. The results are filtered by Standard. The first few rows include:

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
4127719	236713006	X-linked recessive nephrolithiasis with renal failure	Condition	Standard	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4170813	274401005	Uric acid renal calculus	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
4077801	275893001	Phosphate kidney stone	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
4057138	168041003	O/E: renal calculus	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
4057139	168042005	O/E: oxalate renal calculus	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
201620	95570007	Kidney stone	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
4032451	236708007	Calyceal renal calculus	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
4144000	427649000	Calcium renal calculus	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Concept Sets

- Exclude = term is an exclusion from a set built from other terms
- Descendants = include all of this term's descendants as well
- Mapped = include nonstandard terms that map to this term
 - *You don't usually want to do this for Atlas; it may be useful when exporting*

The screenshot shows a software interface titled "Concept Set" with a sub-header "CB - Atlas training pneumonia". The main area displays a table of 7 entries, each representing a concept with its ID, code, name, domain, standard status, and various checkboxes for exclude, descendants, and mapped status. The table includes columns for Concept Id, Concept Code, Concept Name, Domain, Standard Concept Caption, Exclude, Descendants, and Mapped. A legend at the bottom right indicates that orange icons represent Classification, red icons represent Non-Standard, and blue icons represent Standard.

Concept Id	Concept Code	Concept Name	Domain	Standard Concept Caption	Exclude	Descendants	Mapped
261326	75570004	Viral pneumonia	Condition	Standard	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
258180	430395005	Pneumonia due to Gram negative bacteria	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
37116366	733051000	Pneumonia caused by Gram positive bacteria	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
255848	233604007	Pneumonia	Condition	Standard	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4133224	278516003	Lobar pneumonia	Condition	Standard	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
443410	312342009	Infective pneumonia	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
257315	53084003	Bacterial pneumonia	Condition	Standard	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Concept Sets

- Explore results as concept set is built
- See which concepts contribute the most data
- See what nonstandard source terms could have contributed data

The screenshot shows a software interface titled "Concept Set" with the sub-header "CB - Atlas training pneumonia". The main area displays a table of "Included Concepts" with 110 entries. The columns include: Id, Code, Name, Class, RC, DRC, Domain, Vocabulary, and Ancestors. The table lists various types of pneumonia, such as "Infective pneumonia", "Bacterial pneumonia", "Lobar pneumonia", and several types of pneumonia due to different bacteria. The "Vocabulary" column indicates the source of each concept, mostly SNOMED. The "Ancestors" column shows the number of direct ancestors for each concept, ranging from 0 to 3.

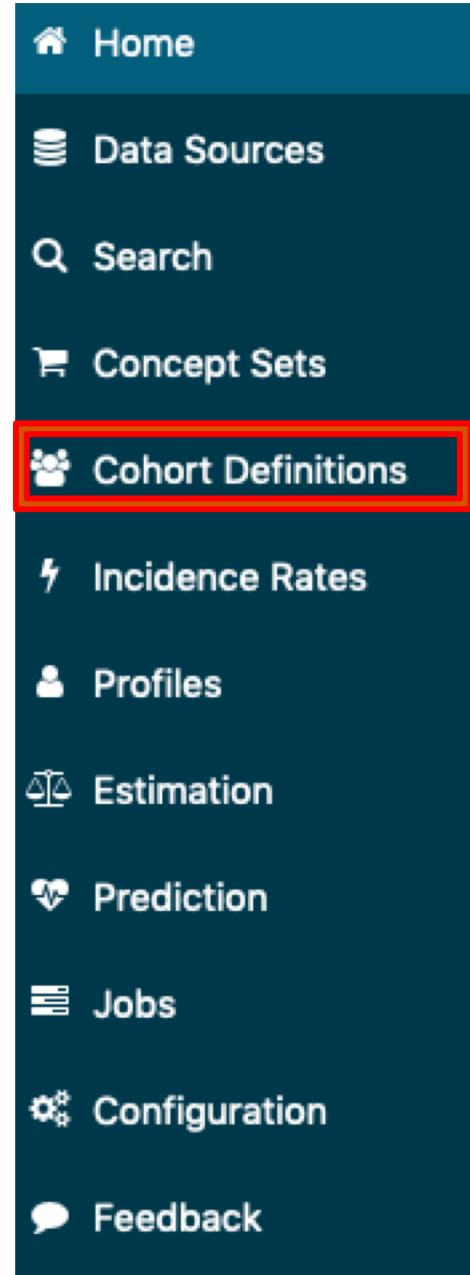
Concept Set								
CB - Atlas training pneumonia								
Concept Set Expression		Included Concepts 110		Included Source Codes		Export		Optimize
Column visibility	Copy	CSV	Show 15	15 entries		Filter:		
Showing 1 to 15 of 110 entries								
Vocabulary								
SNOMED (110)	443410	312342009	Infective pneumonia	Clinical Finding	119,600	298,355	Condition	SNOMED 0
Class								
Clinical Finding (110)	257315	53084003	Bacterial pneumonia	Clinical Finding	61,578	112,481	Condition	SNOMED 0
Condition (110)	4133224	278516003	Lobar pneumonia	Clinical Finding	8,331	31,712	Condition	SNOMED 0
Standard Concept								
Standard (110)	258180	430395005	Pneumonia due to Gram negative bacteria	Clinical Finding	3,004	31,289	Condition	SNOMED 0
Invalid Reason								
Valid (110)	37116366	733051000	Pneumonia caused by Gram positive bacteria	Clinical Finding	0	18,090	Condition	SNOMED 0
Has Records								
false (57)	440431	46970008	Mycoplasma pneumonia	Clinical Finding	15,958	15,958	Condition	SNOMED 3
true (53)	261324	34020007	Pneumonia due to Streptococcus	Clinical Finding	2,144	9,709	Condition	SNOMED 3
Has Descendant Records								
true (55)	259852	22754005	Staphylococcal pneumonia	Clinical Finding	740	8,358	Condition	SNOMED 3
false (55)	4256236	409665004	Pneumonia due to aerobic bacteria	Clinical Finding	27	7,336	Condition	SNOMED 2

Concept Set Tips

- The database contains *standard* terms
- Start with more general terms and use descendants
 - You can optimize a concept set when you're done to do this retrospectively
- Use the “Exclude” option to prune subtrees that you don’t want
- You’ll often want to include Descendants
- Always review included concepts
- Prefix global concept sets with your initials

Cohorts

- Atlas' name for a group of persons meeting a set of rule-based criteria
- Individual rules can include AND and OR logic
- Sequence of rules applied using AND logic (attrition)
- Basis for Atlas' interactions with persons
 - Cohort characterization/descriptives
 - Population-level analyses

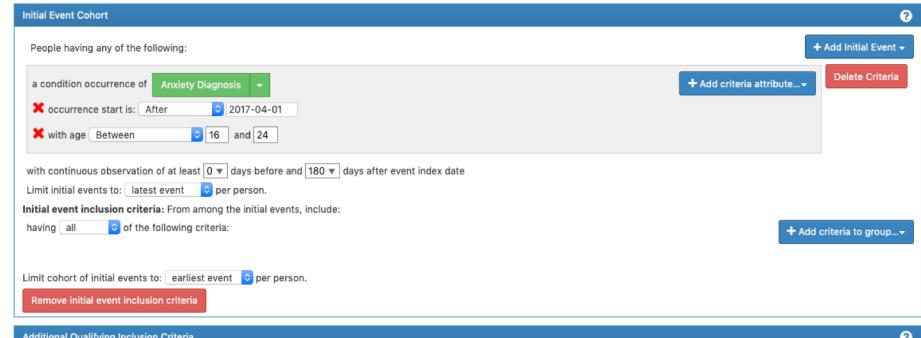


Cohort Etiquette

- All cohorts are global – prefix with initials
- Don't change someone else's cohort (unless they agree)
- OK to pick up concept sets or rules from others' cohorts

Initial Event

- Starting point for cohort attrition
 - Has to be clinical facts
 - Typically includes content and time constraints (*e.g.* persons with diagnosis X plus follow-up for Y months)
-
- Tips
 - Aim for reasonable starting point
 - Don't try to apply all the criteria in this step



The screenshot shows the 'Initial Event Cohort' builder interface. At the top right are buttons for '+ Add Initial Event' (blue), '+ Add criteria attribute...' (blue), and 'Delete Criteria' (red). The main area has a heading 'Initial Event Cohort' and a sub-instruction 'People having any of the following:'. Below this is a search bar with dropdown menus for 'a condition occurrence of' (set to 'Anxiety Diagnosis') and 'occurrence start is' (set to 'After 2017-04-01'). There are also filters for 'with age' (set to 'Between 16 and 24'). To the right of these filters are buttons for '+ Add criteria attribute...' (blue) and 'Delete Criteria' (red). Below the search bar, text specifies 'with continuous observation of at least 0 days before and 180 days after event index date' and 'Limit initial events to: latest event per person.' A section titled 'Initial event inclusion criteria:' allows selecting 'having all' or 'any' of the following criteria. A button '+ Add criteria to group...' is located to the right. At the bottom, there's a note 'Limit cohort of initial events to: earliest event per person.' and a red 'Remove initial event inclusion criteria' button. A link 'Additional Qualifying Inclusion Criteria' is at the very bottom.

Initial Event Inclusion Criteria

- Other things related to the initial event
- Goal is to get an “entry” population, to start the attrition process
- Tips
 - This is where you can select demographics
 - Don’t try to apply all the criteria in this step
 - Huge initial cohort = slow processing

The screenshot shows the 'Initial Event Cohort' configuration page. At the top, there's a header with buttons for '+ Add Initial Event', '+ Add criteria attribute...', and 'Delete Criteria'. Below the header, it says 'People having any of the following:' followed by a dropdown menu set to 'Any Condition'. There are sections for 'with continuous observation of at least [0] days before and [365] days after event index date' and 'Limit initial events to: earliest event [checkbox] per person'. A section titled 'Initial event inclusion criteria' allows selecting multiple conditions from a dropdown menu. Another section for 'with the following event criteria' includes a dropdown menu set to 'FEMALE' with options for 'Add' and 'Import'. At the bottom, there's a 'Limit cohort of initial events to: [all events] [checkbox] per person' section and a 'Remove initial event inclusion criteria' button.

Additional Inclusion Criteria

- Rules that “whittle down” the initial group to what’s needed to answer the question
 - Concurrent inclusion requirements
 - Absence of confounding exclusion requirements

The screenshot shows a software interface for defining inclusion criteria. At the top, there's a green button labeled 'New qualifying inclusion criteria' and a blue header bar with tabs for 'vWD Testing' and 'vWD Diagnosis'. Below the header, there's a section for entering a rule description: 'enter an inclusion rule description' and 'having any of the following criteria:'. A large green box contains the actual rule definition: 'with at least 1 using all occurrences of: a measurement of vWD Testing starting between All days Before and All days After event index date and ending any time, restrict to the same visit occurrence'. There are buttons for 'Copy', 'Delete', '+ Add criteria to group...', '+ Add criteria attribute...', and 'Delete Criteria'.

- Tips
 - Remember AND logic
 - Each rule should accomplish one thing

Inclusion Event Criteria

Name of Cohort

Cohort #348

Cystic Fibrosis

Definition Concept Sets Generation Reporting Export

children with cystic fibrosis

Initial Event Cohort

People having any of the following:

a condition occurrence of Cystic Fibrosis ▾

+ Add Initial Event ▾

Delete Criteria

Cohort index criteria with time constraints

+ Add criteria attribute... ▾

with continuous observation of at least 0 ▾ days before and 548 ▾ days after event index date

Limit initial events to: earliest event ▾ per person.

Add initial event inclusion criteria

The screenshot shows the PEDSnet Cohort builder interface. At the top, it displays 'Cohort #348' and 'Cystic Fibrosis'. Below this is a navigation bar with tabs: Definition, Concept Sets, Generation, Reporting, and Export. The main area is titled 'Initial Event Cohort' and contains a sub-section titled 'People having any of the following:'. Inside this section, there is a green button labeled 'Cystic Fibrosis' with a dropdown arrow. To the right of this button is a blue button labeled '+ Add Initial Event ▾' and a red button labeled 'Delete Criteria'. Above the 'Initial Event Cohort' section, the text 'Name of Cohort' is displayed with an orange arrow pointing to the 'Cystic Fibrosis' button. Below the 'Initial Event Cohort' section, the text 'Cohort index criteria with time constraints' is displayed with an orange arrow pointing to the 'Add criteria attribute...' button. The bottom of the interface has a green button labeled 'Add initial event inclusion criteria'.

Additional Qualifying Inclusion Criteria

- Additional Qualifying Criteria:**
- Second CF diagnosis
 - Drug exclusions

Additional Qualifying Inclusion Criteria ?

New qualifying inclusion criteria

Drug Exclusions Copy Delete

exclude patients with at least one drug exposure of ivacaftor or any of its descendants

having all of the following criteria: + Add criteria to group...

2. Drug Exclusions + Add criteria attribute... Delete Criteria

exclude patients with at least one drug exposure of ivacaftor or any of its descendants

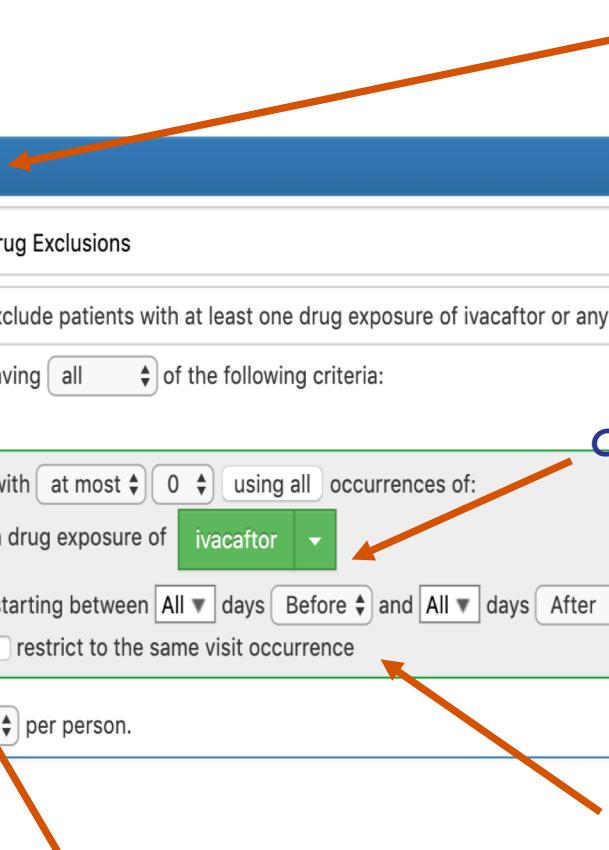
3. Unnamed Criteria

Limit qualifying cohort to: all events per person.

Concept Set

with at most 0 using all occurrences of:
a drug exposure of **ivacaftor** ▼

starting between All days Before and All days After event index date and ending any time.
 restrict to the same visit occurrence



Description
Time constraints – ever on ivacaftor

Best practice for cohort criteria

- Start with largest practical initial cohort
- Keep individual criteria specific, to follow attrition
- Be careful about relative timing of criteria
- Order criteria by desired attrition pattern



Cohort-specific Concept Sets

Cohort #348
Cystic Fibrosis

Definition Concept Sets Generation Reporting Export

New Concept Set Import Export All Concept Sets To CSV

Show 10 entries Filter Cohort Concept Sets:

Id	Title
0	Cystic Fibrosis
1	ivacaftor

Showing 1 to 2 of 2 entries Previous 1 Next

Concept Set Expression Included Concepts 54 Included Source Codes Export Import

Column visibility Copy CSV Show 15 entries Filter:

Showing 1 to 15 of 54 entries Previous 1 2 3 4 Next

Vocabulary	Id	Code	Name	Class	RC	DRC	Domain	Vocabulary	Ancestors
RxNorm (54)	46275587	1655934	ivacaftor 125 MG / lumacaftor 200 MG Oral Tablet [ORKAMBI]	Branded Drug	7,130	7,130	Drug	RxNorm	0
Class	46275586	1655931	ivacaftor / lumacaftor Oral Tablet [ORKAMBI]	Branded Drug Form	0	8,600	Drug	RxNorm	0
Clinical Dose Group (8)	46275585	1655930	ivacaftor 125 MG / lumacaftor 200 MG [ORKAMBI]	Branded Drug Comp	0	7,130	Drug	RxNorm	0
Clinical Drug (8)	46275584	1655929	ORKAMBI	Brand Name	0	0	Drug	RxNorm	0
Branded Drug (7)	46275583	1655928	ivacaftor 125 MG / lumacaftor 200 MG Oral Tablet	Clinical Drug	6,472	13,602	Drug	RxNorm	0
Branded Dose Group (6)	46275582	1655926	ivacaftor / lumacaftor Oral Tablet	Clinical Drug Form	0	16,401	Drug	RxNorm	0
Branded Drug Comp (6)	46275579	1655921	ivacaftor 125 MG	Clinical Drug Comp	0	16,443	Drug	RxNorm	0
Domain	46221757	1606870	ivacaftor 75 MG Oral Granules [Kalydeco]	Branded Drug	335	335	Drug	RxNorm	0
Drug (54)									
Standard Concept									
Standard (38)									
Classification (14)									
Non-Standard (2)									

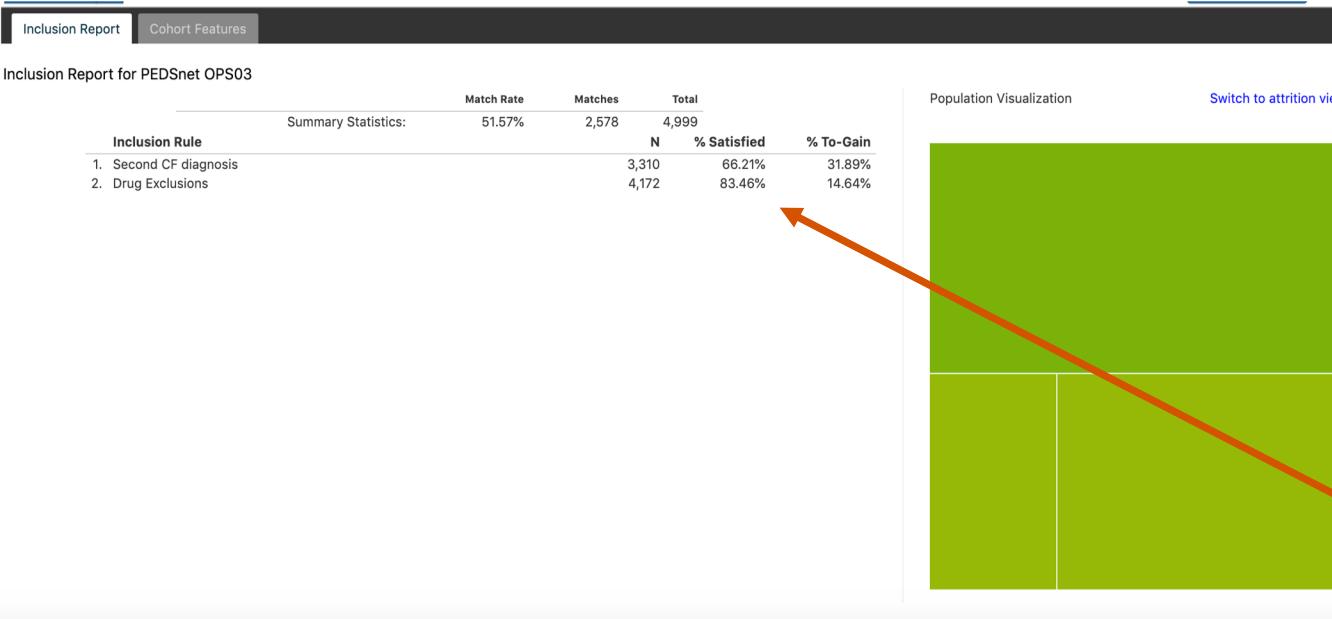
ivacaftor codeset

Generating cohorts

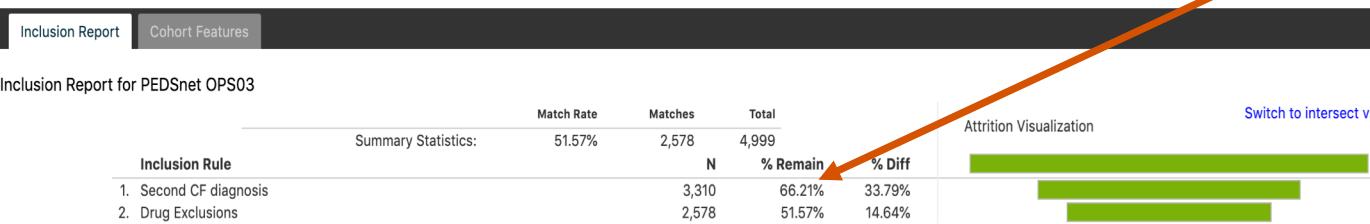
- Just Generate first to test attrition
 - Time depends on cohort complexity and size
 - Usually minutes
- Subsequently can Generate With Features to better examine the cohort
 - Takes much longer – expect ~1 hour for substantial cohort

Generating Cohort: Attrition

- Provides line-by-line report for each inclusion criteria



- 4,999 patients who meet initial criteria
- 33.79% have a second diagnosis of CF
- 14.64% have a drug exclusion
- Shows each individual criteria
- Shows attrition view



Generating Cohort Features

- Cohort Features provides distributions for the generated cohort
- Provides demographic, condition, drug, procedures, measurements, observation distributions from data model
- Also provides Diabetes Comorbidity Severity Index (DCSI), Charlson Index, CHADS2VASc, CHADS2 distributions for cohort

Atlas Features: Demographic Distributions

Demographic features

Feature Report for PEDSnet OPS03
Features are baseline characteristics (e.g collected before /on cohort start)

Demographics Conditions Drugs Procedures Measurements Observations Distributions

Column visibility Copy CSV Show 15 entries Filter: Previous 1 2 3 Next

Showing 1 to 15 of 38 entries

Name	Count	% of cohort
age group: 00-04	926	36.00
age group: 05-09	435	16.90
age group: 10-14	474	18.40
age group: 15-19	484	18.80
age group: 20-24	128	5.00
age group: 25-29	50	2.00
age group: 30-34	27	1.10
age group: 35-39	14	0.60
age group: 40-44	18	0.70
Hispanic or Latino	163	6.40
Not Hispanic or Latino	2,242	87.00
FEMALE	1,275	49.50
MALE	1,303	50.60
index month: 1	1,138	44.20
index month: 10	135	5.30

Showing 1 to 15 of 38 entries Previous 1 2 3 Next

Analysis
Index Month (12)
Index Year (9)
Age Group (9)
Race (4)
Gender (?)
Time Window
None (38)

Atlas Features: Conditions Distributions

Definitions and lookback period

Long Term: 365 day lookback. Medium Term: 180d lookback Short Term: 30d lookback. Overlapping: Event spans cohort start date.

Column visibility Copy CSV Show 15 entries

Showing 1 to 15 of 940 entries

	Concept Name	Time Window	Person Count	% of cohort
<input type="button" value="Explore"/>	Cystic fibrosis	Long Term	2,495	96.80
<input type="button" value="Explore"/>	Inherited mucociliary clearance defect	Long Term	2,495	96.80
<input type="button" value="Explore"/>	Mucociliary clearance defect	Long Term	2,495	96.80
<input type="button" value="Explore"/>	Inherited mucociliary clearance defect	Short Term	2,066	80.20
<input type="button" value="Explore"/>	Mucociliary clearance defect	Short Term	2,066	80.20
<input type="button" value="Explore"/>	Cystic fibrosis	Short Term	2,066	80.20
<input type="button" value="Explore"/>	Cystic fibrosis without meconium ileus	Long Term	905	35.20
<input type="button" value="Explore"/>	Cystic fibrosis of the lung	Long Term	834	32.40
<input type="button" value="Explore"/>	Inflammation of specific body systems	Long Term	734	28.50
<input type="button" value="Explore"/>	Inflammatory disorder of the respiratory system	Long Term	622	24.20
<input type="button" value="Explore"/>	Inflammatory disorder of the respiratory tract	Long Term	621	24.10
<input type="button" value="Explore"/>	Ear, nose and throat disorder	Long Term	609	23.70
<input type="button" value="Explore"/>	Cystic fibrosis without meconium ileus	Short Term	591	23.00
<input type="button" value="Explore"/>	Inflammation of specific body organs	Long Term	569	22.10

Sorting by different fields

Filter: Previous 1 2 3 4 5 ... 63 Next

Can explore all descendants of a concept

The screenshot shows a table of 15 rows, each representing a medical condition or symptom. The columns are: Concept Name, Time Window, Person Count, and % of cohort. The 'Time Window' column is annotated with a blue arrow pointing to its header. The 'Explore' button in the first column of each row is also annotated with a blue arrow. A third blue arrow points to the sorting/filtering controls at the top right of the table.

	Concept Name	Time Window	Person Count	% of cohort
<input type="button" value="Explore"/>	Cystic fibrosis	Long Term	2,495	96.80
<input type="button" value="Explore"/>	Inherited mucociliary clearance defect	Long Term	2,495	96.80
<input type="button" value="Explore"/>	Mucociliary clearance defect	Long Term	2,495	96.80
<input type="button" value="Explore"/>	Inherited mucociliary clearance defect	Short Term	2,066	80.20
<input type="button" value="Explore"/>	Mucociliary clearance defect	Short Term	2,066	80.20
<input type="button" value="Explore"/>	Cystic fibrosis	Short Term	2,066	80.20
<input type="button" value="Explore"/>	Cystic fibrosis without meconium ileus	Long Term	905	35.20
<input type="button" value="Explore"/>	Cystic fibrosis of the lung	Long Term	834	32.40
<input type="button" value="Explore"/>	Inflammation of specific body systems	Long Term	734	28.50
<input type="button" value="Explore"/>	Inflammatory disorder of the respiratory system	Long Term	622	24.20
<input type="button" value="Explore"/>	Inflammatory disorder of the respiratory tract	Long Term	621	24.10
<input type="button" value="Explore"/>	Ear, nose and throat disorder	Long Term	609	23.70
<input type="button" value="Explore"/>	Cystic fibrosis without meconium ileus	Short Term	591	23.00
<input type="button" value="Explore"/>	Inflammation of specific body organs	Long Term	569	22.10

Exporting and Importing

- Atlas allows for easy sharing of cohorts across multiple instances of data model
- As user builds cohort, backend JSON and SQL constructed that makes sharing possible
- Backend JSON / SQL can be copied to another instance and regenerated in another database without needing to manually recreate all individual criteria
- SQL generator available for MSSQL, Oracle, Postgres, Amazon Red Shift, Impala, Netezza

Exporting and Importing

Cohort #348
Cystic Fibrosis

Definition Concept Sets Generation Reporting Export

Text View Graphical View JSON SQL

Template OHDSI.SQL MSSQL Server MS APS Oracle PostgreSQL Amazon Red Shift Impala Netezza

[Copy To Clipboard](#)

```
CREATE TABLE #Codesets (
    codeset_id int NOT NULL,
    concept_id bigint NOT NULL
)
;

INSERT INTO #Codesets (codeset_id, concept_id)
SELECT 0 as codeset_id, c.concept_id FROM (select distinct I.concept_id FROM
(
    select concept_id from @vocabulary_database_schema.CONCEPT where concept_id in (44782927,44802278,42538542,4127073,44805713,42538541,441267,4341770,254320,44802
) I
) C;
INSERT INTO #Codesets (codeset_id, concept_id)
SELECT 1 as codeset_id, c.concept_id FROM (select distinct I.concept_id FROM
(
    select concept_id from @vocabulary_database_schema.CONCEPT where concept_id in (42705869,36248963,36238290,36238291,46275584,35200179,36248598,36248599,4270932
) I
) C;

with primary_events (event_id, person_id, start_date, end_date, op_start_date, op_end_date, visit_occurrence_id) as
(
-- Begin Primary Events
select P.ordinal as event_id, P.person_id, P.start_date, P.end_date, op_start_date, op_end_date, cast(P.visit_occurrence_id as bigint) as visit_occurrence_id
FROM
```

“Exporting” Reports

The screenshot shows the PEDSnet Cohort builder interface. The title bar says "Cohort #16" and the cohort name is "vsd_surgeries". The top navigation bar includes tabs for "Definition", "Concept Sets", "Generation", "Reporting", and "Export". Below the tabs are buttons for "Text View" (selected), "Graphical View", "JSON", and "SQL". A "Copy To Clipboard" button is also present. The main content area displays the cohort definition:

```
vsd_surgeries

Initial Event Cohort
People having any of the following:
• a condition occurrence of VSD3
  ◦ with age <= 1

with continuous observation of at least 0 days prior and 0 days after event index date, and limit initial events to: all events per person.

For people matching the Primary Events, include:
Having all of the following criteria:
• at least 1 occurrences of a condition occurrence of VSD3
  starting between all days Before and 1 days Before event index date
• and at least 2 occurrences of a procedure of Echo2
  starting between all days Before and 0 days After event index date
• and exactly 0 occurrences of a procedure of CV surgery1
  starting between all days Before and 1 days Before event index date

Limit cohort of initial events to: earliest event per person.

Inclusion Rules
Inclusion Criteria #1: has CV surgery in the next 2 years
Having all of the following criteria:
• at least 1 occurrences of a procedure of CV surgery1
  starting between 1 days After and all days After event index date

Limit qualifying cohort to: earliest event per person.
```

vsd_surgeries

Initial Event Cohort

People having any of the following:

- a condition occurrence of VSD³
 - with age <= 1

with continuous observation of at least 0 days prior and 0 days after event index date, and limit initial events to: **all events per person**.

For people matching the Primary Events, include:

Having all of the following criteria:

- at least 1 occurrences of a condition occurrence of VSD³
 - starting between all days Before and 1 days Before event index date
- and at least 2 occurrences of a procedure of Echo²
 - starting between all days Before and 0 days After event index date
- and exactly 0 occurrences of a procedure of CV surgery¹
 - starting between all days Before and 1 days Before event index date

Limit cohort of initial events to: **earliest event per person**.

Inclusion Rules

Inclusion Criteria #1: has CV surgery in the next 2 years

Having all of the following criteria:

- at least 1 occurrences of a procedure of CV surgery¹
 - starting between 1 days After and all days After event index date

Limit qualifying cohort to: **earliest event per person**.

“Exporting” Reports

Cohort #16

vsd_surgeries

Definition Concept Sets Generation Reporting Export

Text View Graphical View JSON SQL

Primary Criteria

Results will be generated for every person event matching the following primary criterion. Final results will be limited to the first events matching any additional criteria and inclusion rules. Result index dates will be the start dates of the matching primary criteria events.

All of condition: VSD

A horizontal timeline from -1.0 to 1.0. A blue dot is placed exactly at 0.0, indicating the start date of the matching primary criteria event.

Additional Criteria

Restrict to people having events matching all of the following criteria. Events must start within bracketed period () relative to index date. Lines and arrows represent required duration of these events.

All of condition: VSD

and procedure: Echo

and procedure: CV surgery

The timeline shows three events. The first event, 'VSD', is a solid blue dot at index 0 labeled 'it 1 occurrence'. The second event, 'Echo', is a dashed orange line with an arrow pointing right, labeled 'At least 2 occurrences'. The third event, 'CV surgery', is a solid blue dot at index 0 labeled 'occurrences'.

Inclusion Rules

has CV surgery in the next 2 years

All of procedure: CV surgery

The timeline shows a single event starting at index 0, indicated by a blue dot. An orange vertical bar extends to the right from index 0, labeled 'At least 1 occu'.

Restrict to people having events matching all of the following criteria. Events must start within bracketed period () relative to index date. Lines and arrows represent required duration of these events.

The timeline shows a single event starting at index 0, indicated by a blue dot. An orange vertical bar extends to the right from index 0, labeled 'At least 1 occu'.

Atlas Cohort Walk-Throughs

Atlas Example 1: Type 2 Diabetes

Feature	Definition
Inclusion	<ul style="list-style-type: none">• 0-17 years old• Has diagnoses T2DM• Does not have T1DM
Exclusion	<ul style="list-style-type: none">• DKA, liver disease, history of malignancy within 5 years, major surgery, amputation, renal disease, history of hereditary glucose-galactose malabsorption, hypoglycemia, T2DM drugs, use of corticosteroids, anticonvulsant history, pregnancy, alcohol or drug abuse
Other Features	<ul style="list-style-type: none">• Metformin and insulin acceptable for use• Need to examine how presence of T1D affects the cohort