

Introduction to Capr

Martin Lavallee

<https://github.com/OHDSI/Capr>

Capr 1.0.3

Reference

Articles ▾

Changelog





Capr

Capr is part of [HADES](#).

Cohort definition Application Programming in R

Introduction

Capr is an R package to develop and manipulate OHDSI cohort definitions. This package assists in creating a cohort definition that can be compiled by circe-be using CirceR. Cohort definitions developed in Capr are compatible with OHDSI ATLAS. Additionally the package allows for development of cohort design components, sub-items of a cohort design that are meant to be reusable and mutable to assist creating cohorts in study development.

System Requirements

Requires R (version 3.6.0 or higher). Installation on Windows requires [RTools](#). Libraries used in Capr require Java. Capr requires a connection to an OMOP vocabulary database to query concepts.

Links

- [Browse source code](#)
- [Report a bug](#)
- [Ask a question](#)

License

Apache License 2.0

Citation

[Citing Capr](#)

Developers

Martin Lavallee
Author, maintainer

Dev status

 R-CMD-check passing

 codecov 59%

What is Capr?



ATLAS

Home Data Sources Search Concept Sets Cohort Definitions Characterizations Cohort Pathways Incidence Rates Profiles Estimation Prediction Jobs Configuration Feedback

Cohort #1770676
created by anonymous on 2019-06-03 11:33

[BookOfOHDSt] New users of Thiazide-like diuretics as first-line monotherapy for hypertension

Definition Concept Sets Generation Samples Reporting Export Versions Messages 2

Enter a cohort definition description here

Cohort Entry Events

Events having any of the following criteria:

a drug exposure of Thiazide or thiazide-like diuretic
✗ for the first time in the person's history

with continuous observation of at least 365 days before and 0 days after event index date

Limit initial events to: earliest event per person.

Restrict initial events

Inclusion Criteria

New inclusion criteria

1. has hypertension diagnosis in 1 yr prior to treatment
2. Has no prior antihypertensive drug exposures in medical history
3. Is only taking ACE as monotherapy, with no concomitant combination treatments

Limit qualifying events to: earliest event per person.

```
#Capr Cohort Definition |
cse <- Capr::getConceptIdDetails(conceptIds = c(907013, 978555, 974166, 1395058),
                                connectionDetails = connectionDetails,
                                vocabularyDatabaseSchema = vocabularyDatabaseSchema) %>%
  createConceptSetExpression(Name = "Thiazide")

# create primary criteria -----
#create drug exposure query for first line hypertension drugs
drug <- createDrugExposure(conceptSetExpression = cse)
#build pc
pc <- createPrimaryCriteria(Name = "primary criteria",
                           ComponentList = list(drug),
                           ObservationWindow = createObservationWindow(),
                           Limit = "All")

es <- createCustomEraEndStrategy(ConceptSetExpression = cse, gapDays = 30, offset = 0)
cer <- createCohortEra(EraPadDays = 30L)

cd <- createCohortDefinition(Name = "Thiazide Use Cohort",
                             PrimaryCriteria = pc,
                             EndStrategy = es,
                             CohortEra = cer)
```

Key: ATLAS JSON = Capr JSON

Why use Capr?

Don't Repeat Yourself (DRY) Principal

Scenario

- 50 Adv
- 3 varia
 - Bas
 - Inpa
 - Age

Less code matters
Less code matters
Less code matters
Less code matters
Less code matters
Less code matters
Less code matters
Less code matters
Less code matters
Less code matters
Less code matters
Less code matters

rules
concept
nostics



Cohort Templates

ATLAS Cohort #1779557
created by anonymous on 2022-06-14 20:55

[caprDemo]-sglt2_t2d

Definition Concept Sets Generation Samples Reporting Export Versions Messages 1

Enter a cohort definition description here

Cohort Entry Events

Events having any of the following criteria:

a drug exposure of SGLT2 + Add attribute...

✗ for the first time in the person's history Delete Criteria

✗ occurrence start is: after 2012-01-01 + Add attribute...

✗ with age Greater or Equal To 18 Delete Criteria

+ Add Initial Event...

with continuous observation of at least 365 days before and 0 days after event index date

Limit initial events to: all events per person.

Restrict initial events to:

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

with exactly 0 using all occurrences of:

a condition occurrence of Type 1 Diabetes Diagnosis + Add attribute...

where event starts between

All days Before and 0 days After index start date [add additional constraint](#)

The index date refers to the event from the Cohort Entry criteria.

☐ restrict to the same visit occurrence

☐ allow events from outside observation period Delete Criteria

Limit initial events to: earliest event per person.

Remove initial event restriction

Inclusion Criteria

New inclusion criteria

1. Case for T2DM using algorithm Keep these pieces

Limit qualifying events to: earliest event per person.

Apache 2.0 open source software provided by OHDSI join the journey

Cohort Templates

```
1 #Template function of drug exposure with T2D
2 drugExposureWithT2D <- function(connectionDetails,
3                                 vocabularyDatabaseSchema,
4                                 drugClassConceptIds,
5                                 drugClassName,
6                                 irs,
7                                 ac) {
8   #Step 1: lookup drug -----
9   drug <- Capr::getConceptIdDetails(
10     conceptIds = drugClassConceptIds,
11     connectionDetails = connectionDetails,
12     vocabularyDatabaseSchema = vocabularyDatabaseSchema)
13   #Step 2: Turn into CSE -----
14   drugCSE <- Capr::createConceptSetExpression(
15     conceptSet = drug,
16     Name = drugClassName,
17     includeDescendants = TRUE
18   )
19   #Step 3: Create Drug Exposure Query -----
20   drugQuery <- Capr::createDrugExposure(conceptSetExpression = drug,
21                                         attributeList = list(
22                                           Capr::createAgeAttribute(Op = "gte", Value = 18),
23                                           Capr::createFirstAttribute(),
24                                           Capr::createOccurrenceStartDateAttribute(Op = "gt",
25                                                                 Value = "2012-01-01")
26                                         ))
27   #Step 4: Create Primary Criteria -----
28   ow <- Capr::createObservationWindow(PriorDays = 365, PostDays = 0)
29   pc <- Capr::createPrimaryCriteria(Name = paste(drugClassName, "Exposure"),
30                                    ComponentList = list(drugQuery),
31                                    ObservationWindow = ow,
32                                    Limit = "All")
33   #Step 5: Add other pieces of cohort -----
34   #Note IRS and AC saved
35   # Create exit strategy
36   es <- Capr::createCustomEraEndStrategy(ConceptSetExpression = drug,
37                                          gapDays = 45L,
38                                          offset = 0)
39   # Censoring events
40   cen <- Capr::createCensoringCriteria(Name = "Death Censoring Criteria",
41                                       ComponentList = list(Capr::createDeath()))
42   #Cohort Era
43   cohortEra <- createCohortEra(EraPadDays = 30L)
44   # Step 6: Create cohort definition -----
45   cd <- Capr::createCohortDefinition(Name = "New SGLT2 Users for T2D Cohort",
46                                     PrimaryCriteria = pc,
47                                     AdditionalCriteria = ac,
48                                     InclusionRules = irs,
49                                     EndStrategy = es,
50                                     CensoringCriteria = cen,
51                                     CohortEra = cohortEra)
52   return(cd)
53 }
54 }
```

```
1 # Step 0: Set up Analysis -----
2
3 #load libraries
4 library(DatabaseConnector)
5 library(Capr)
6 library(CohortGenerator)
7 library(CohortDiagnostics)
8 library(FeatureExtraction)
9
10 #load Capr template function
11 source("drugExposureWithT2D.R")
12
13 #create connection details
14 connectionDetails <- DatabaseConnector::createConnectionDetails(
15   dbms = "postgresql",
16   server = keyring::key_get("cdm_server"),
17   user = keyring::key_get("cdm_user"),
18   password = keyring::key_get("cdm_password"),
19   port = "5441"
20 )
21 # set vocabulary database schema
22 vocabularyDatabaseSchema <- "cdm"
23 cdmDatabaseSchema <- "cdm"
24 resultsDatabaseSchema <- "results"
25
26 # Step 1: Identify drug concepts -----
27 # Create drug concept list using OMOP ids
28 drugConceptsList <- list(
29   'sglt2' = c(793293, 35198021, 35198029, 43009020, 43526465, 44785829, 45774751),
30   'glp1' = c(793143, 40170911, 43013171, 44816332, 45774435, 1583722),
31   'ddp4' = c(1580747, 19122137, 35197921, 35198118, 40166035, 40239216,
32             43008991, 43009051, 43009070, 43009089, 43013884),
33   'sulfonylureas' = c(1502809, 1502855, 1530014, 1559684, 1560171, 1594973,
34                      1597756, 19001409, 19033498, 19059796, 19097821, 40798860),
35   'metformin' = c(1503297)
36 )
37
38 # Step 2: Load inclusion rules and additional criteria -----
39 irs <- Capr::loadComponent("cohorts/components/phekbT2dCase.json")
40 ac <- Capr::loadComponent("cohorts/components/noT1D_AC.json")
41
42 # Step 4: Build Cohort Definitions -----
43
44 set <- purrr::map2(drugConceptsList,
45                   names(drugConceptsList),
46                   ~drugExposureWithT2D(connectionDetails = connectionDetails,
47                                         vocabularyDatabaseSchema = vocabularyDatabaseSchema,
48                                         drugClassConceptIds = .x,
49                                         drugClassName = .y,
50                                         irs = irs,
51                                         ac = ac)) %>%
52   Capr::createCohortDataframe()
53
54 # Step 5: Generate Cohorts -----
55 CohortGenerator::getCohortTableNames()
56
57 CohortGenerator::generateCohortSet(connectionDetails = connectionDetails,
58                                   cdmDatabaseSchema = cdmDatabaseSchema,
59                                   cohortDatabaseSchema = resultsDatabaseSchema,
60                                   cohortTableNames = CohortGenerator::getCohortTableNames(),
61                                   cohortDefinitionSet = set)
```

Long Capr
Script or
scrape JSON

Brief Tutorial (part 1)

Step 1: Create Concept Set Expressions

```
#Type 2 Diabetes Medications CSE
T2Rx <- c(1502809L, 1502826L, 1503297L, 1510202L,
         1515249L, 1516766L, 1525215L, 1529331L,
         1530014L, 1547504L, 1559684L, 1560171L,
         1580747L, 1583722L, 1594973L, 1597756L) %>%
  getConceptIdDetails(
    connectionDetails = connectionDetails,
    vocabularyDatabaseSchema = vocabularyDatabaseSchema
  ) %>%
  createConceptSetExpression(
    Name = "Type 2 Diabetes Medications",
    includeDescendants = TRUE
  )
```

Step 2: Create Query

```
# T2Rx Query
T2RxQuery <- createDrugExposure(conceptSetExpression = T2Rx)

#HbA1c Query with value attribute
AbLabHbA1cQuery <- createMeasurement(
  conceptSetExpression = AbLabHbA1c,
  attributeList = list(
    #add attribute of >= 6.5 %
    createValueAsNumberAttribute(
      op = "gte",
      value = 6.5)
  )
)
```

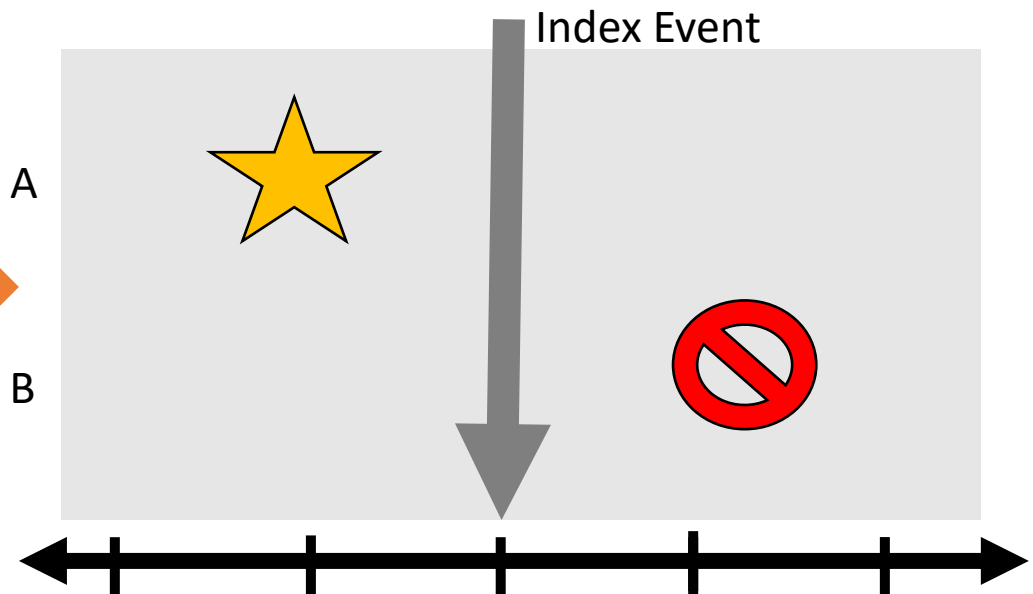
Step 3: Create Count

```
t11 <- createTimeline(startwindow = createwindow(
  StartDays = "All", StartCoeff = "Before",
  EndDays = 0L, EndCoeff = "After"))

atLeast1T2RxCount <- createCount(Query = T2RxQuery,
  Logic = "at_least",
  Count = 1L,
  Timeline = t11)
```

Patient A

Patient B



Brief Tutorial (part 2)

Step 4: Create Groups

```
atLeast1AbLabGroup <- createGroup(  
  Name = "Abnormal labs for HbA1c, Fasting+Random Glucose",  
  type = "ANY",  
  criteriaList = list(  
    atLeast1AbLabHbA1cCount,  
    atLeast1AbLabFastingGlucCount,  
    atLeast1AbLabRandomGlucCount)  
)  
  
#4) Path 1: 0 T2Dx, 1+ T2Rx and 1+ AbLab  
Pathway1T2DMGroup <- createGroup(  
  Name = "Pathway1",  
  Description = "0 T2Dx, 1+ T2Rx and 1+ AbLab",  
  type = "ALL",  
  criteriaList = list(notT2DxCount, atLeast1T2RxCount),  
  Groups = list(atLeast1AbLabGroup))
```

Step 6: Create Cohort Definition

```
#Create Cohort Definition  
T2DMPhenotype <- createCohortDefinition(  
  Name = "PheKB T2DM Definition",  
  PrimaryCriteria = PrimaryCriteria,  
  AdditionalCriteria = AdditionalCriteria,  
  InclusionRules = InclusionRules,  
  CensoringCriteria = CensoringCriteria  
)  
  
### compile circe  
T2DMPhenotypeJson <- compileCohortDefinition(T2DMPhenotype)
```

Step 5: Create Cohort Pieces

```
#create Inclusion Rules  
#keep T2DM cases if they meet 1 of the 5 pathways  
  
T2DMCase <- createGroup(  
  Name = "Case for T2DM using algorithm",  
  type = "ANY",  
  Groups = list(Pathway1T2DMGroup, Pathway2T2DMGroup,  
    Pathway3T2DMGroup, Pathway4T2DMGroup,  
    Pathway5T2DMGroup)  
)  
  
InclusionRules <- createInclusionRules(  
  Name = "IRs for T2DM Case Phenotype",  
  Contents = list(T2DMCase),  
  Limit = "First"  
)  
  
Capr::saveComponent(InclusionRules,  
  saveName = "phekbT2dCase",  
  savePath = "cohorts/components")
```

Brief Tutorial (part 3)

Start with JSON and edit
using Capr

```
#import json
sglt2Cohort <- Capr::readInCirc(jsonPath = "cohorts/json/sglt2.json",
                              connectionDetails = connectionDetails,
                              vocabularyDatabaseSchema = vocabularyDatabaseSchema)

#lookup drug
glp1 <- Capr::getConceptIdDetails(
  conceptIds = c(793143, 40170911, 43013171, 44816332, 45774435, 1583722),
  connectionDetails = connectionDetails,
  vocabularyDatabaseSchema = vocabularyDatabaseSchema)
# Turn into CSE
glp1CSE <- Capr::createConceptSetExpression(
  conceptSet = glp1,
  Name = "GLP1",
  includeDescendants = TRUE
)
#Create Drug Exposure Query
glp1Query <- Capr::createDrugExposure(conceptSetExpression = glp1CSE,
                                     attributeList = list(
                                       Capr::createAgeAttribute(Op = "gte", value = 18),
                                       Capr::createFirstAttribute(),
                                       Capr::createOccurrenceStartDateAttribute(Op = "gt",
                                                                                   value = "2012-01-01")
                                     ))
#Create Primary Criteria
ow <- Capr::createObservationWindow(PriorDays = 365, PostDays = 0)
pc <- Capr::createPrimaryCriteria(Name = "GLP1 Exposure",
                                 ComponentList = list(glp1Query),
                                 ObservationWindow = ow,
                                 Limit = "All")

glp1Cohort <- sglt2Cohort
glp1Cohort@PrimaryCriteria <- pc
```


Thank you!

Please explore Capr and provide feedback <https://github.com/OHDSI/Capr/>