Introduction to Capr

Martin Lavallee

https://github.com/OHDSI/Capr

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 3.5.0 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.
What is Capr?

Key: ATLAS JSON = Capr JSON
Why use Capr?

Don’t Repeat Yourself (DRY) Principal

Scenario 1
• 50 Adverse Drug Reactions
• 3 variations of the cohorts
  • Base
  • Inpatient Visit
  • Age > 17

Scenario 2
• Prostate Cancer
• 10 variations of inclusion rules
• Need to update a single concept ID after each cohort diagnostics
Cohort Templates

Need to change this:
DDP4, GLP1, metformin, Sulfonylureas, etc

Keep these pieces
Cohort Templates

Long Capr Script or scrape JSON
Brief Tutorial (part 1)

Step 1: Create Concept Set Expressions

```r
# Type 2 Diabetes Medications CSE
T2RX <- c(1502809L, 1502828L, 1503297L, 1510202L,
          1515249L, 1516768L, 1525215L, 1529331L,
          1530014L, 1547504L, 1559084L, 1560171L,
          1580747L, 1583722L, 1594973L, 1597756L)

getConceptDetails(
  connectionDetails = connectionDetails,
  vocabularyDatabaseSchema = vocabularyDatabaseSchema
)

createConceptSetExpression(
  Name = "Type 2 Diabetes Medications",
  includeDescendants = TRUE
)
```

Step 2: Create Query

```r
# T2RX Query
T2RXquery <- createDrugExposure(conceptSetExpression = T2RX)

# HbA1c Query with value attribute
A1cHbALquery <- createMeasurement(
  conceptSetExpression = A1cHbALc,
  attributeList = list(
    addAttributeOf = "HbA1c",
    value = 6.5
  )
)
```

Step 3: Create Count

```r
t11 <- createTimeline(StartWindow = createWindow(
  StartDays = "All", StartCoeff = "Before",
  EndDays = 0L, EndCoeff = "After")
)

atLeast1T2RXCount <- createCount(Query = T2RXquery,
  Logic = "at_least",
  count = 1L,
  Timeline = t11)
```

Index Event

Patient A

Patient B
Brief Tutorial (part 2)

Step 4: Create Groups

```r
atleast1Ablabgroup <- createGroup(
  name = "Abnormal labs for HBA1c, Fasting+Random Glucose",
  type = "ANY",
  criteriaList = list(
    atleast1AblabHba1cCount,
    atleast1AblabFastingGlucCount,
    atleast1AblabRandomGlucCount
  )
)

#4) Path 1: 0 T2Dx, 1+ T2R and 1+ Ablab
Pathway1T2DMGroup <- createGroup(
  name = "Pathway1",
  description = "0 T2Dx, 1+ T2R and 1+ Ablab",
  type = "ALL",
  criteriaList = list(noT2DxCount, atleast1T2RxCOUNT),
  groups = list(atleast1AblabGroup)
)
```

Step 5: Create Cohort Pieces

```r
#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")
```

Step 6: Create Cohort Definition

```r
#Create Cohort Definition
T2DMPhenotype <- createCohortDefinition(
  name = "PheKB T2DM Definition",
  primaryCriteria = PrimaryCriteria,
  additionalCriteria = AdditionalCriteria,
  inclusionRules = InclusionRules,
  censoringCriteria = censoringCriteria
)

## compile circe
T2DMPhenotypeJson <- compileCohortDefinition(T2DMPhenotype)
```
Start with JSON and edit using Capr

```r
# import json
sigt2cohort <- capr::readncohrce(cohorts/json/sglt2.json,
                                connectionDetails = connectionDetails,
                                vocabularydatabaseschema = vocabularydatabaseschema)

# lookup drug
glp1 <- Capr::getConceptIdDetails(
  ConceptIds = c(793143, 40170031, 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::createDrugExposureConceptSetExpression(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 <- sigt2cohort

glp1Cohort@PrimaryCriteria <- pc
```
Thank you!

Please explore Capr and provide feedback [https://github.com/OHDSI/Capr/](https://github.com/OHDSI/Capr/)