Designing a study using the TreatmentPatterns package

Aniek F Markus

2021-07-16

This vignette describes how to design a custom treatment patterns study.

1. Setting up

Make sure you have TreatmentPatterns installed and all dependencies (see instructions in README).

2. Designing the study

There are two approaches: A) manually filling the inst folder in the package or B) by using function arguments when calling the main function of the package. We first describe both approaches for databases in OMOP CDM format (2.1) and then for other data formats (2.2).

Throughout this vignette we will use an adjusted example describing the treatment patterns of hypertension patients. The included drugs of interests are hydrochlorothiazide, metorolol, amlodipine, lisinopril, and losartan (for simplicity we limit it to 5 events of interest).

Some definitions used throughout this vignette:

- Target cohort = study population.

- Event cohort(s) = treatment(s) of interest.

2.1 Databases in OMOP CDM format

A) Manually filling the inst folder in the package

1. Define target/event cohorts.

- Specify all cohorts in ATLAS. When defining the target cohort note that it might be desirable to request a minimum follow up time after cohort entry / the index date to have sufficient information on treatment history.

- Download the SQL and JSON files and add to folder inst/cohorts/SQL and inst/cohorts/JSON.

- Then specify the target/event cohorts in inst/settings/cohorts_to_create.csv:

cohortId	cohortName	cohortDefinition	cohortType	atlasID
Unique ID number	Descriptive name cohort	'ATLAS' or 'Custom' (see explanation in Section 3.1)	'target' or 'event'	Cohort ID ATLAS

2. Optional: specify baseline characteristics of interest.

covariateName	covariateId
Descriptive name covariate	Unique ID number referring to covariate from FeatureExtraction or 'Custom' (see explanation in Section 3.2)

3. Define study settings.

A default list of study settings:

param	values	description
studyName	default	Unique name identifying the set of study parameters below
targetCohortId	1	Select one study population
eventCohortIds	"10,11,12,13,14"	Select all treatments of interest
includeTreatmentsPriorToIndex	0	Number of days prior to the index date of the target cohort that event cohorts are allowed to start
minEraDuration	0	Minimum time an event era should last to be included in analysis
splitEventCohorts		Specify event cohort to split in acute (< 30 days) and therapy ($>= 30$ days)
eraCollapseSize	0	Window of time between which two eras of the same event cohort are collapsed into one era
combinationWindow	30	Window of time two event cohorts need to overlap to be considered a combination treatment
minStepDuration	30	Minimum time an event era before or after a generated combination treatment should last to be included in analysis
filterTreatments	First	Select first occurrences of / changes between / all event cohorts
maxPathLength	5	Maximum number of steps included in treatment pathway
minCellCount	0	Minimum number of persons with a specific treatment pathway for the pathway to be included in analysis
minCellMethod	Remove	Select to completely remove / sequentially adjust (by removing last step as often as necessary) treatment pathways below minCellCount
groupCombinations	10	Select to group all non-fixed combinations in one category 'other' in the sunburst plot
addNoPaths	FALSE	Select to include untreated persons without treatment pathway in the sunburst plot

Change these parameters according to the needs of your study in inst/Settings/study_settings.csv.

B) Using function arguments when calling the main function of the package

We need the following components:

1. Define target/event cohorts.

- Specify all cohorts in ATLAS. When defining the target cohort note that it might be desirable to request a minimum follow up time after cohort entry / the index date to have sufficient information on treatment history.

- The study population of interest (target cohorts).

```
targetCohorts <- data.frame(cohortId = c(1),
                            atlasId = c(1777380),
                            cohortName = c('Hypertension'))
```

- The events of interest (event cohorts).

```
eventCohorts <- data.frame(cohortId = c(10, 11, 12, 13, 14),
                           atlasId = c(1777381,1777382, 1777383, 1777384, 1777385),
                           cohortName = c('Hydrochlorothiazide', 'Metorolol',
                                          'Amlodipine', 'Lisinopril', 'Losartan'))
```

2. Optional: specify baseline characteristics of interest.

```
characterizationSettings <- data.frame(covariateName = c('Male', 'Age',
                                                          'Charlson comorbidity index score'),
                                        covariateId = c(8507001, 1002, 1901))
```

3. Define study settings.

```
studySettings <- data.frame(studyName = c("default", "analysis1"),
                            targetCohortId = c(1,1),
                            eventCohortIds = c("10,11,12,13,14","10,11,12,13,14"),
                            includeTreatmentsPriorToIndex = c(0,0),
                            minEraDuration = c(0,5),
                            splitEventCohorts = c("",""),
                            eraCollapseSize = c(0,30),
                            combinationWindow = c(30,30),
                            minStepDuration = c(30,30),
                            filterTreatments = c("First","Changes"),
                            maxPathLength = c(5,5),
                            minCellCount = c(0,5),
                            minCellMethod = c("Remove","Adjust"),
                            groupCombinations = c(10,10),
                            addNoPaths = c(FALSE, TRUE))
```

2.2 Databases in other formats

A) Manually filling the inst folder in the package

1. Define target/event cohorts and add to package.

- Import the target/event cohorts from a csv file into the package. Add cohorts in inst/cohorts/input_cohorts.csv:

cohortId	personId	startDate	endDate
Unique ID number	Unique person ID number	Entry date cohort	Exit date cohort

- Then specify the target/event cohorts in inst/settings/cohorts_to_create.csv:

cohortId	cohortName	cohortDefinition	cohortType	atlasID
Unique ID number	Descriptive name cohort	'ATLAS' or 'Custom' (see explanation in Section 3.1)	'target' or 'event'	Cohort ID ATLAS

2. Define study settings.

See explanation in Section 2.1 (A).

B) Using function arguments when calling the main function of the package

1. Define target/event cohorts and add to package.

- Import the target/event cohorts from a csv file into the package. Add cohorts in inst/cohorts/input_cohorts.csv:

cohortId	personId	startDate	endDate
Unique ID number	Unique person ID number	Entry date cohort	Exit date cohort

- The study population of interest (target cohorts).

```
targetCohorts <- data.frame(cohortId = c(1),
                            atlasId = c(""),
                            cohortName = c('Hypertension'))
```

- The events of interest (event cohorts).

```
eventCohorts <- data.frame(cohortId = c(10, 11, 12, 13, 14),
                           atlasId = c("","", "", "", ""),
                           cohortName = c('Hydrochlorothiazide', 'Metorolol',
                                          'Amlodipine', 'Lisinopril', 'Losartan'))
```

2. Define study settings.

See explanation in Section 2.1 (B).

3. Extra options

3.1 Custom cohorts using concept set and template

Instead of defining each cohort separately in ATLAS, there is the possibility to specify a file with concept sets that will be used in combination with a SQL template. In the package there is a cohort template available

for drugs (inst/SQL/CohortDrugTemplate.sql). This template identifies all drug exposures of the concept set till the end of a continuous exposure with a maximum persistence window of 30 days. If you want to make use of this option, you need to fill the inst folder manually (approach A).

- Alternatively, specify a concept set and create re-use or create a new inst/Settings/CohortTemplate.sql. Specify concept sets in inst/Settings/eventcohorts_custom.csv:

cohortName	count	conceptSet
Descriptive name cohort	Number of concept IDs in conceptSet	List of unique concept IDs to be included

3.2 Adding custom covariates for baseline characterization

If desired, one can add custom covariates in inst/SQL for characterization. An example template:

```
SELECT @covariateId AS covariate_id,
{@aggregated} ? {
  cohort_definition_id,
  COUNT(DISTINCT target.@rowIdField) AS sum_value
} : {
  target.@rowIdField  AS row_id,
  1 AS covariate_value
}
FROM @cohortTable target
INNER JOIN @cdmDatabaseSchema.condition_occurrence covariate
ON covariate.person_id = target.@rowIdField
WHERE covariate.condition_concept_id IN ()
AND covariate.condition_start_date <= target.cohort_start_date
{@cohortId != -1} ? {AND cohort_definition_id IN (@cohortId)}
{@aggregated} ? {GROUP BY cohort_definition_id}
```

3.3 Stand alone sunburst plot functionality

```
data <- readr::read_csv(paste0(system.file(package = "TreatmentPatterns"),"/examples/sequences.csv"),
                        col_types = list("c", "i"))
createSunburstPlot(data, folder = outputFolder <- paste0(getwd(), "/output"),
                   file_name = "example", shiny = FALSE,
                   title = "example sequences.csv")
```

3.4 Add custom analysis parts

If desired, one can add additional output functions. Need to add R code and adjust shiny application.