Package ‘CohortSymmetry’

July 26, 2024

Title  Sequence Symmetry Analysis Using the Observational Medical Outcomes Partnership Common Data Model

Version  0.1.2

Maintainer  Xihang Chen <xihang.chen@ndorms.ox.ac.uk>

Description  Calculating crude sequence ratio, adjusted sequence ratio and confidence intervals using data mapped to the Observational Medical Outcomes Partnership Common Data Model.

License  Apache License (>= 2)

Encoding  UTF-8

RoxygenNote  7.3.1

Suggests  testthat (>= 3.1.5), DBI (>= 1.0.0), cli, checkmate, odbc, RPostgres, tidyselect, knitr, dbplyr (>= 2.5.0), rmarkdown

Imports  CDMConnector (>= 1.3.0), dplyr, ggplot2, magrittr, PatientProfiles, rlang, stringr, tibble, visOmopResults (>= 0.3.0), tidyr, omopgenerics (>= 0.2.1), flextable, gt, DrugUtilisation (>= 0.5.0), CodelistGenerator (>= 3.1.0)

Config/testthat/edition  3

URL  https://oxford-pharmacoepi.github.io/CohortSymmetry/

BugReports  https://github.com/oxford-pharmacoepi/CohortSymmetry/issues

VignetteBuilder  knitr

NeedsCompilation  no

Author  Xihang Chen [aut, cre] (<https://orcid.org/0009-0001-8112-8959>), Tyman Stanford [aut] (<https://orcid.org/0000-0002-8570-5493>), Berta Raventós [aut] (<https://orcid.org/0000-0002-4668-2970>), Nicole Pratt [aut] (<https://orcid.org/0000-0001-8730-8910>), Ed Burn [aut] (<https://orcid.org/0000-0002-9286-1128>), Marti Català [aut] (<https://orcid.org/0000-0003-3308-9905>), Danielle Newby [aut] (<https://orcid.org/0000-0002-3001-1478>), Núria Mercadé-Besora [aut] (<https://orcid.org/0009-0006-7948-3747>), Mike Du [aut] (<https://orcid.org/0000-0002-9517-8834>),
generateSequenceCohortSet

Intersecting the index and marker cohorts prior to calculating Sequence Symmetry Ratios

Description

Join two tables in the CDM (one for index and the other for marker cohorts) into a new table in the cdm taking into account the maximum time interval between events. Index and marker cohorts should be instantiated in advance by the user.

Usage

generateSequenceCohortSet(
  cdm,
  indexTable,
  markerTable,
  name,
  cohortDateRange = as.Date(c(NA, NA)),
  indexId = NULL,
  markerId = NULL,
  daysPriorObservation = 0,
  washoutWindow = 0,
  indexMarkerGap = NULL,
  combinationWindow = c(0, 365)
)
Arguments

cdm A CDM reference.
indexTable A table in the CDM that the index cohorts should come from.
markerTable A table in the CDM that the marker cohorts should come from.
nname The name within the cdm that the output is called. Default is joined_cohorts.
cohortDateRange Two dates indicating study period and the sequences that the user wants to restrict to.
indexId Cohort definition IDs in indexTable to be considered for the analysis. Change to NULL if all indices are wished to be included.
markerId Cohort definition IDs in markerTable to be considered for the analysis. Change to NULL if all markers are wished to be included.
daysPriorObservation The minimum amount of prior observation required on both the index and marker cohorts per person.
washoutWindow A washout window to be applied on both the index cohort event and marker cohort.
indexMarkerGap The maximum allowable gap between the end of the first episode and the start of the second episode in a sequence/combination.
combinationWindow A constrain to be placed on the gap between two initiations. Default c(0,365), meaning the gap should be larger than 0 but less than or equal to 365.

Value

A table within the cdm reference.

Examples

library(CohortSymmetry)
cdm <- mockCohortSymmetry()
cdm <- generateSequenceCohortSet(
  cdm = cdm,
  name = "joined_cohorts",
  indexTable = "cohort_1",
  markerTable = "cohort_2"
)
cdm$joined_cohorts
CDMCconnector::cdmDisconnect(cdm = cdm)
mockCohortSymmetry  

Description

Creates a mock cdm with two default synthetic cohorts, one is the index cohort and the other one is the marker cohort. However the users could specify them should they wish.

Usage

mockCohortSymmetry(
  seed = 1,
  indexCohort = NULL,
  markerCohort = NULL,
  con = DBI::dbConnect(duckdb::duckdb(), "memory:"),
  schema = "main"
)

Arguments

- **seed**: The seed to be inputted.
- **indexCohort**: The tibble of your index cohort. Default is NULL, which means the default indexCohort is being used.
- **markerCohort**: The tibble of your marker cohort. Default is NULL, which means the default markerCohort is being used.
- **con**: Connection detail.
- **schema**: Name of your write schema.

Value

A mock cdm object contains your index and marker cohort

Examples

library(CohortSymmetry)
cdm <- mockCohortSymmetry()
cdm
CDMConnector::cdmDisconnect(cdm = cdm)
**plotSequenceRatios**

A plot for the sequence ratios.

**Description**

It provides a ggplot of the sequence ratios of index and marker cohorts.

**Usage**

```r
plotSequenceRatios(
  result,
  onlyaSR = FALSE,
  plotTitle = NULL,
  labs = c("SR", "Drug Pairs"),
  colours = c("red", "blue")
)
```

**Arguments**

- `result`: Table output from `summariseSequenceRatios`.
- `onlyaSR`: If the only SR to be plotted is the adjusted SR.
- `plotTitle`: Title of the plot, if NULL no title will be plotted.
- `labs`: Axis labels for the plot.
- `colours`: Colours for both parts of the plot, pre- and post- time 0.

**Value**

A plot for the sequence ratios of index and marker cohorts.

**Examples**

```r
library(CohortSymmetry)

cdm <- mockCohortSymmetry()
cdm <- generateSequenceCohortSet(cdm = cdm,
  indexTable = "cohort_1",
  markerTable = "cohort_2",
  name = "joined_cohort")

sequence_ratio <- summariseSequenceRatios(cohort = cdm$joined_cohort,
  minCellCount = 0)

plotSequenceRatios(result = sequence_ratio)
CDMConnector::cdmDisconnect(cdm = cdm)
```
plotTemporalSymmetry

A plot for the temporal symmetry of cohorts.

Description

It provides a ggplot of the temporal symmetry of two or more cohorts.

Usage

```
plotTemporalSymmetry(
  result,
  plotTitle = NULL,
  labs = c("Time (months)", "Individuals (N)"),
  xlim = c(-12, 12),
  colours = c("blue", "red"),
  scales = "free"
)
```

Arguments

- `result`: Table output from summariseTemporalSymmetry.
- `plotTitle`: Title of the plot, if NULL no title will be plotted.
- `labs`: Axis labels for the plot.
- `xlim`: Limits for the x axis of the plot.
- `colours`: Colours for both parts of the plot, pre- and post- time 0.
- `scales`: Whether to set free y scales for the facet wrap when there are multiple plots (i.e. each plot has its own scaled y axis) or set them equal for all. Only accepts "free" for the former and "fixed" for the latter.

Value

A plot for the temporal symmetry of cohorts.

Examples

```
library(CohortSymmetry)
cdm <- mockCohortSymmetry()
cdm <- generateSequenceCohortSet(cdm = cdm,
  indexTable = "cohort_1",
  markerTable = "cohort_2",
  name = "joined_cohort")
temporal_symmetry <- summariseTemporalSymmetry(cohort = cdm$joined_cohort,
  minCellCount = 0)
plotTemporalSymmetry(result = temporal_symmetry)
CDMConnector::cdmDisconnect(cdm = cdm)
```
**summariseSequenceRatios**

*Sequence ratio calculations*

**Description**

Using `generateSequenceCohortSet` to obtain sequence ratios for the desired outcomes.

**Usage**

```r
summariseSequenceRatios(
  cohort,
  cohortId = NULL,
  confidenceInterval = 95,
  movingAverageRestriction = 548,
  minCellCount = 5
)
```

**Arguments**

- **cohort**: A cohort table in the cdm.
- **cohortId**: The Ids in the cohort that are to be included in the analyses.
- **confidenceInterval**: Default is 95, indicating the central 95% confidence interval.
- **movingAverageRestriction**: The moving window when calculating nSR, default is 548.
- **minCellCount**: The minimum number of events to reported, below which results will be obscured. If 0, all results will be reported.

**Value**

A local table with all the analyses.

**Examples**

```r
library(CohortSymmetry)
cdm <- mockCohortSymmetry()
cdm <- generateSequenceCohortSet(cdm = cdm,
  name = "joined_cohorts",
  indexTable = "cohort_1",
  markerTable = "cohort_2")
pssa_result <- summariseSequenceRatios(cohort = cdm$joined_cohorts)
pssa_result
CDMConnector::cdmDisconnect(cdm)
```
**summariseTemporalSymmetry**

_Summarise temporal symmetry_

---

**Description**

Using `generateSequenceCohortSet` to obtain temporal symmetry (aggregated counts) of two cohorts.

**Usage**

```r
summariseTemporalSymmetry(
  cohort,
  cohortId = NULL,
  timescale = "month",
  minCellCount = 5
)
```

**Arguments**

- **cohort**: A cohort table in the cdm.
- **cohortId**: The Ids in the cohort that are to be included in the analyses.
- **timescale**: Timescale for the x axis of the plot (month, day, year).
- **minCellCount**: The minimum number of events to be reported, below which results will be obscured. If 0, all results will be reported.

**Value**

An aggregated table with difference in time (marker - index) and the relevant counts.

**Examples**

```r
library(CohortSymmetry)
cdm <- mockCohortSymmetry()
cdm <- generateSequenceCohortSet(cdm = cdm,
  name = "joined_cohorts",
  indexTable = "cohort_1",
  markerTable = "cohort_2")
temporal_symmetry <- summariseTemporalSymmetry(cohort = cdm$joined_cohorts)
```
The `tableSequenceRatios` function provides a formatted visualization of `sequence_symmetry` objects. It returns a formatted version of the `sequence_symmetry` object.

**Description**

It provides a formatted table with the contents of the `summariseSequenceRatios` output.

**Usage**

```r
tableSequenceRatios(
  result,
  type = "gt",
  estimateNameFormat = c(\`N (%)\` = "<count> (<percentage> %)", \`SR (CI)\` = "<point_estimate> (<lower_CI> - <upper_CI>)"),
  style = "default",
  studyPopulation = TRUE,
  cdmName = TRUE,
  .options = NULL
)
```

**Arguments**

- **result**: A `sequence_symmetry` object.
- **type**: Type of desired formatted table, possibilities: "gt", "flextable", "tibble".
- **estimateNameFormat**: The columns that the user wishes to see for the formatted table, by default it would display both the counts and sequence ratios.
- **style**: Named list that specifies how to style the different parts of a gt table or flextable. See `visOmopResults` package for more information on how to define a style. Alternatively, use "default" to get `visOmopResults` style, or NULL for gt/flextable default styling.
- **studyPopulation**: whether to report the study population.
- **cdmName**: whether to report database names.
- **.options**: named list with additional formatting options. `tableSequenceRatiosOptions()` shows allowed arguments and their default values.

**Value**

A formatted version of the `sequence_symmetry` object.
Examples

```r
library(CohortSymmetry)
cdm <- mockCohortSymmetry()
cdm <- generateSequenceCohortSet(cdm = cdm,
    indexTable = "cohort_1",
    markerTable = "cohort_2",
    name = "joined_cohort")
res <- summariseSequenceRatios(cohort = cdm$joined_cohort)
gtResult <- tableSequenceRatios(res)
CDMConnector::cdmDisconnect(cdm = cdm)
```

---

tableSequenceRatiosOptions

*A formatted visualization of sequence.ratios objects.*

Description

It provides a list of allowed inputs for `.option` argument in `tableSequenceRatios` and their given default value.

Usage

```r
tableSequenceRatiosOptions()
```

Value

The default `.options` named list.

Examples

```r
{
  tableSequenceRatiosOptions()
}
```
Index

generateSequenceCohortSet, 2
mockCohortSymmetry, 4
plotSequenceRatios, 5
plotTemporalSymmetry, 6
summariseSequenceRatios, 7
summariseTemporalSymmetry, 8
tableSequenceRatios, 9
tableSequenceRatiosOptions, 10