Package 'QuantifyingBiasInApapStudies'

October 21, 2019

Type Package

Title Quantifying Bias in Epidemiological Studies on the Association Between Acetaminophen and Cancer

Version 1.0.0

Date 2019-10-21

Maintainer Martijn Schuemie <schuemie@ohdsi.org>

Description Quantifying Bias in Epidemiological Studies on the Association Between Acetaminophen and Cancer.

License Apache License 2.0

Depends DatabaseConnector

Imports SqlRender,
CaseControl (>= 2.0.0),
CohortMethod,
ParallelLogger,
Cyclops,
FeatureExtraction,
EmpiricalCalibration,
ff,
rmarkdown,
ggplot2,
gridExtra,
scales

NeedsCompilation no

RoxygenNote 6.1.1

Encoding UTF-8

R topics documented:

`createCcAnalysesDetails` *Create the case-control analyses details*

Description

Create the case-control analyses details

Usage

```
createCcAnalysesDetails(workFolder)
```

Arguments

`workFolder`	Name of local folder to place results; make sure to use forward slashes (/)

Details

This function creates files specifying the case-control analyses that will be performed.

`createCmAnalysesDetails` *Create the analyses details*

Description

Create the analyses details

Usage

```
createCmAnalysesDetails(workFolder)
```

Arguments

`workFolder`	Name of local folder to place results; make sure to use forward slashes (/)

Details

This function creates files specifying the analyses that will be performed.

createCohorts *Create the exposure and outcome cohorts*

Description

Create the exposure and outcome cohorts

Usage

```
createCohorts(connectionDetails, cdmDatabaseSchema, cohortDatabaseSchema,
  cohortTable = "cohort", oracleTempSchema, outputFolder)
```

Arguments

connectionDetails	An object of type connectionDetails as created using the createConnectionDetails function in the DatabaseConnector package.
cdmDatabaseSchema	Schema name where your patient-level data in OMOP CDM format resides. Note that for SQL Server, this should include both the database and schema name, for example 'cdm_data.dbo'.
cohortDatabaseSchema	Schema name where intermediate data can be stored. You will need to have write priviliges in this schema. Note that for SQL Server, this should include both the database and schema name, for example 'cdm_data.dbo'.
cohortTable	The name of the table that will be created in the work database schema. This table will hold the exposure and outcome cohorts used in this study.
oracleTempSchema	Should be used in Oracle to specify a schema where the user has write priviliges for storing temporary tables.
outputFolder	Name of local folder to place results; make sure to use forward slashes (/)

Details

This function will create the exposure and outcome cohorts following the definitions included in this package.

createCustomCovariatesSettings *Create custom covariate settings*

Description

Create custom covariate settings

Usage

```
createCustomCovariatesSettings(useBmi = FALSE, useAlcohol = FALSE,
  useSmoking = FALSE, useDiabetesMedication = FALSE,
  useRheumatoidArthritis = FALSE, useNonRa = FALSE,
  useFatigue = FALSE, useMigraine = FALSE)
```

Arguments

useBmi	Create a covariate for BMI (prior to cohort start).
useAlcohol	Create a covariate for alcohol use (prior to cohort start).
useSmoking	Create a covariate for smoking (prior to cohort start).
useDiabetesMedication	Create a covariate for diabetes (medication use) (prior to cohort start).
useRheumatoidArthritis	Create a covariate for RA (prior to cohort start).
useNonRa	Create a covariate for non-RA, chronic back or chronic neck pain (prior to cohort start).
useFatigue	Create a covariate for fatigue or lack of energy (prior to cohort start).
useMigraine	Create a covariate for migraine or chronic headache (prior to cohort start).

---

createPlotsAndTables *Generate plots and tables*

Description

Generate plots and tables

Usage

```
createPlotsAndTables(connectionDetails, cdmDatabaseSchema,
  oracleTempSchema, outputFolder, blind = TRUE)
```

Arguments

connectionDetails	An object of type connectionDetails as created using the createConnectionDetails function in the DatabaseConnector package.
cdmDatabaseSchema	Schema name where your patient-level data in OMOP CDM format resides. Note that for SQL Server, this should include both the database and schema name, for example 'cdm_data.dbo'.
oracleTempSchema	Should be used in Oracle to specify a schema where the user has write priviliges for storing temporary tables.
outputFolder	Name of local folder where the results were generated; make sure to use forward slashes (/). Do not use a folder on a network drive since this greatly impacts performance.
blind	Blind results? If true, no real effect sizes will be shown. To be used during development.

Details

Requires that the CohortMethod and CaseControl analyses have been executed

---

`execute` *Execute the Study*

Description

Execute the Study

Usage

```
execute(connectionDetails, cdmDatabaseSchema,
  cohortDatabaseSchema = cdmDatabaseSchema, cohortTable = "cohort",
  oracleTempSchema = cohortDatabaseSchema, outputFolder,
  createCohorts = TRUE, runCohortMethod = TRUE,
  runCaseControl = TRUE, createPlotsAndTables = TRUE,
  generateReport = TRUE, maxCores = 4, blind = TRUE)
```

Arguments

`connectionDetails`	An object of type `connectionDetails` as created using the `createConnectionDetails` function in the DatabaseConnector package.
`cdmDatabaseSchema`	Schema name where your patient-level data in OMOP CDM format resides. Note that for SQL Server, this should include both the database and schema name, for example 'cdm_data.dbo'.
`cohortDatabaseSchema`	Schema name where intermediate data can be stored. You will need to have write priviliges in this schema. Note that for SQL Server, this should include both the database and schema name, for example 'cdm_data.dbo'.
`cohortTable`	The name of the table that will be created in the work database schema. This table will hold the exposure and outcome cohorts used in this study.
`oracleTempSchema`	Should be used in Oracle to specify a schema where the user has write priviliges for storing temporary tables.
`outputFolder`	Name of local folder to place results; make sure to use forward slashes (/). Do not use a folder on a network drive since this greatly impacts performance.
`createCohorts`	Create the cohortTable table with the exposure and outcome cohorts?
`runCohortMethod`	Perform the cohort method analyses? Requires the cohorts have been created.
`runCaseControl`	Perform the case-control analyses? Requires the cohorts have been created.
`createPlotsAndTables`	Generate output plots and tables? Requires CohortMethodd and CaseControl analyses have been completed.
`generateReport`	Generate a report document? Requires the plots and tables have been created.
`maxCores`	How many parallel cores should be used? If more cores are made available this can speed up the analyses.
`blind`	Blind results? If true, no real effect sizes will be shown. To be used during development.

Details

This function executes the QuantifyingBiasInApapStudies Study.

The `createCohorts`, `synthesizePositiveControls`, `runAnalyses`, and `runDiagnostics` arguments are intended to be used to run parts of the full study at a time, but none of the parts are considerd to be optional.

Examples

```
## Not run:
connectionDetails <- createConnectionDetails(dbms = "postgresql",
                                             user = "joe",
                                             password = "secret",
                                             server = "myserver")

execute(connectionDetails,
        cdmDatabaseSchema = "cdm_data",
        cohortDatabaseSchema = "study_results",
        cohortTable = "cohort",
        oracleTempSchema = NULL,
        outputFolder = "c:/temp/study_results",
        maxCores = 4)

## End(Not run)
```

---

generateReport *Generate a report containing the main results*

Description

Generate a report containing the main results

Usage

```
generateReport(outputFolder)
```

Arguments

`outputFolder`	Name of local folder to place results; make sure to use forward slashes (/). Do not use a folder on a network drive since this greatly impacts performance.

`getDbCustomCovariatesData` *Get custom covariate information from the database*

Description

Constructs custom covariates for a cohort.

Usage

```
getDbCustomCovariatesData(connection, oracleTempSchema = NULL,
  cdmDatabaseSchema, cohortTable = "#cohort_person", cohortId = -1,
  cdmVersion = "5", rowIdField = "subject_id", covariateSettings,
  aggregated = FALSE)
```

Arguments

`connection`	A connection to the server containing the schema as created using the `connect` function in the `DatabaseConnector` package.
`oracleTempSchema`	A schema where temp tables can be created in Oracle.
`cdmDatabaseSchema`	The name of the database schema that contains the OMOP CDM instance. Requires read permissions to this database. On SQL Server, this should specifiy both the database and the schema, so for example 'cdm_instance.dbo'.
`cohortTable`	Name of the table holding the cohort for which we want to construct covariates. If it is a temp table, the name should have a hash prefix, e.g. '#temp_table'. If it is a non-temp table, it should include the database schema, e.g. 'cdm_database.cohort'.
`cohortId`	For which cohort ID should covariates be constructed? If set to -1, covariates will be constructed for all cohorts in the specified cohort table.
`cdmVersion`	The version of the Common Data Model used. Currently only `cdmVersion = "5"` is supported.
`rowIdField`	The name of the field in the cohort temp table that is to be used as the row_id field in the output table. This can be especially usefull if there is more than one period per person.
`covariateSettings`	An object of type `covariateSettings` as created using the `createCustomCovariatesSettings` function.
`aggregated`	Should aggregate statistics be computed instead of covariates per cohort entry?

Details

This function uses the data in the CDM to construct a large set of covariates for the provided cohort. The cohort is assumed to be in an existing temp table with these fields: 'subject_id', 'cohort_definition_id', 'cohort_start_date'. Optionally, an extra field can be added containing the unique identifier that will be used as rowID in the output. Typically, users don't call this function directly.

Value

Returns an object of type `covariateData`, containing information on the baseline covariates. Information about multiple outcomes can be captured at once for efficiency reasons. This object is a list with the following components:

covariates An ffdf object listing the baseline covariates per person in the cohorts. This is done using a sparse representation: covariates with a value of 0 are omitted to save space. The covariates object will have three columns: rowId, covariateId, and covariateValue. The rowId is usually equal to the person_id, unless specified otherwise in the rowIdField argument.

covariateRef An ffdf object describing the covariates that have been extracted.

metaData A list of objects with information on how the covariateData object was constructed.

`QuantifyingBiasInApapStudies` *QuantifyingBiasInApapStudies*

Description

QuantifyingBiasInApapStudies

`runCaseControl` *Run CohortMethod package*

Description

Run CohortMethod package

Usage

```
runCaseControl(connectionDetails, cdmDatabaseSchema, cohortDatabaseSchema,
  cohortTable, oracleTempSchema, outputFolder, maxCores)
```

Arguments

`connectionDetails`
: An object of type `connectionDetails` as created using the `createConnectionDetails` function in the DatabaseConnector package.

`cdmDatabaseSchema`
: Schema name where your patient-level data in OMOP CDM format resides. Note that for SQL Server, this should include both the database and schema name, for example 'cdm_data.dbo'.

`cohortDatabaseSchema`
: Schema name where intermediate data can be stored. You will need to have write priviliges in this schema. Note that for SQL Server, this should include both the database and schema name, for example 'cdm_data.dbo'.

`cohortTable`
: The name of the table that will be created in the work database schema. This table will hold the exposure and outcome cohorts used in this study.

`oracleTempSchema`
: Should be used in Oracle to specify a schema where the user has write priviliges for storing temporary tables.

`outputFolder`
: Name of local folder where the results were generated; make sure to use forward slashes (/). Do not use a folder on a network drive since this greatly impacts performance.

`maxCores`
: How many parallel cores should be used? If more cores are made available this can speed up the analyses.

Details

Run the CohortMethod package, which implements the comparative cohort design.

`runCohortMethod` *Run CohortMethod package*

Description

Run CohortMethod package

Usage

```
runCohortMethod(connectionDetails, cdmDatabaseSchema, cohortDatabaseSchema,
  cohortTable, oracleTempSchema, outputFolder, maxCores)
```

Arguments

`connectionDetails`
: An object of type `connectionDetails` as created using the `createConnectionDetails` function in the DatabaseConnector package.

`cdmDatabaseSchema`
: Schema name where your patient-level data in OMOP CDM format resides. Note that for SQL Server, this should include both the database and schema name, for example 'cdm_data.dbo'.

`cohortDatabaseSchema`
: Schema name where intermediate data can be stored. You will need to have write priviliges in this schema. Note that for SQL Server, this should include both the database and schema name, for example 'cdm_data.dbo'.

`cohortTable`
: The name of the table that will be created in the work database schema. This table will hold the exposure and outcome cohorts used in this study.

`oracleTempSchema`
: Should be used in Oracle to specify a schema where the user has write priviliges for storing temporary tables.

`outputFolder`
: Name of local folder where the results were generated; make sure to use forward slashes (/). Do not use a folder on a network drive since this greatly impacts performance.

`maxCores`
: How many parallel cores should be used? If more cores are made available this can speed up the analyses.

Details

Run the CohortMethod package, which implements the comparative cohort design.

Index