Package ‘IbdCharacterization’

January 26, 2022

Title Characterization of Inflammatory Bowel Disease Patient Cohorts

Version 1.0

Description IbdCharacterization describes the baseline demographic and clinical characteristics, as well as the occurrence of treatments and outcomes of individuals diagnosed with inflammatory bowel disease (IBD) and, specifically, with Crohn's disease and ulcerative colitis.

Depends DatabaseConnector (>= 2.2.0)

Imports CohortDiagnostics (>= 1.2.2),
FeatureExtraction,
ParallelLogger,
OhdsiSharing (>= 0.2.1),
SqlRender,
digest,
readr,
dplyr,
ggplot2 (<= 3.2.1),
tibble

Suggests knitr,
rmarkdown

Remotes OHDSI/CohortDiagnostics,
OHDSI/FeatureExtraction,
OHDSI/ParallelLogger,
OHDSI/SqlRender,
OHDSI/OhdsiSharing

License Apache License 2.0

VignetteBuilder knitr

URL `https://github.com/ohdsi-studies/IbdCharacterization`

BugReports `https://github.com/ohdsi-studies/IbdCharacterization/issues`

Encoding UTF-8

LazyData TRUE

RoxygenNote 7.1.2

R topics documented:

`createBulkCharacteristics` *Create cohort characteristics in bulk*

Description

This function will perform the same actions as the main RunStudy.R::runStudy() function but in a single SQL operation.

Usage

```
createBulkCharacteristics(
  connection,
  oracleTempSchema,
  cohortIds,
  cdmDatabaseSchema,
  cohortDatabaseSchema,
  cohortTable
)
```

`createCohortTable` *Create cohort table(s)*

Description

This function creates an empty cohort table. Optionally, additional empty tables are created to store statistics on the various inclusion criteria.

Usage

```
createCohortTable(
  connectionDetails = NULL,
  connection = NULL,
  cohortDatabaseSchema,
  cohortTable = "cohort",
  createInclusionStatsTables = FALSE,
  resultsDatabaseSchema = cohortDatabaseSchema,
  cohortInclusionTable = paste0(cohortTable, "_inclusion"),
  cohortInclusionResultTable = paste0(cohortTable, "_inclusion_result"),
  cohortInclusionStatsTable = paste0(cohortTable, "_inclusion_stats"),
  cohortSummaryStatsTable = paste0(cohortTable, "_summary_stats")
)
```

Arguments

`connectionDetails`
: An object of type `connectionDetails` as created using the `createConnectionDetails` function in the DatabaseConnector package. Can be left NULL if `connection` is provided.

`connection`
: An object of type `connection` as created using the `connect` function in the DatabaseConnector package. Can be left NULL if `connectionDetails` is provided, in which case a new connection will be opened at the start of the function, and closed when the function finishes.

`cohortDatabaseSchema`
: Schema name where your cohort table resides. Note that for SQL Server, this should include both the database and schema name, for example 'scratch.dbo'.

`cohortTable`
: Name of the cohort table.

`createInclusionStatsTables`
: Create the four additional tables for storing inclusion rule statistics?

`resultsDatabaseSchema`
: Schema name where the statistics tables reside. Note that for SQL Server, this should include both the database and schema name, for example 'scratch.dbo'.

`cohortInclusionTable`
: Name of the inclusion table, one of the tables for storing inclusion rule statistics.

`cohortInclusionResultTable`
: Name of the inclusion result table, one of the tables for storing inclusion rule statistics.

`cohortInclusionStatsTable`
: Name of the inclusion stats table, one of the tables for storing inclusion rule statistics.

`cohortSummaryStatsTable`
: Name of the summary stats table, one of the tables for storing inclusion rule statistics.

`getCohortCharacteristics`

Create characterization of a cohort

Description

Computes features using all drugs, conditions, procedures, etc. observed on or prior to the cohort index date.

Usage

```
getCohortCharacteristics(
  connectionDetails = NULL,
  connection = NULL,
  cdmDatabaseSchema,
  oracleTempSchema = NULL,
  cohortDatabaseSchema = cdmDatabaseSchema,
  cohortTable = "cohort",
  cohortId,
  covariateSettings = FeatureExtraction::createDefaultCovariateSettings()
)
```

Arguments

`connectionDetails`
: An object of type `connectionDetails` as created using the `createConnectionDetails` function in the DatabaseConnector package. Can be left NULL if `connection` is provided.

`connection`
: An object of type `connection` as created using the `connect` function in the DatabaseConnector package. Can be left NULL if `connectionDetails` is provided, in which case a new connection will be opened at the start of the function, and closed when the function finishes.

`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 privileges for storing temporary tables.

`cohortDatabaseSchema`
: Schema name where your cohort table resides. Note that for SQL Server, this should include both the database and schema name, for example 'scratch.dbo'.

`cohortTable`
: Name of the cohort table.

`cohortId`
: The cohort definition ID used to reference the cohort in the cohort table.

`covariateSettings`
: Either an object of type `covariateSettings` as created using one of the createCovariate functions in the FeatureExtraction package, or a list of such objects.

Value

A data frame with cohort characteristics.

---

getInclusionStatistics

Get statistics on cohort inclusion criteria

Description

Get statistics on cohort inclusion criteria

Usage

```
getInclusionStatistics(
  connectionDetails = NULL,
  connection = NULL,
  resultsDatabaseSchema,
  cohortId,
  simplify = TRUE,
  cohortTable = "cohort",
  cohortInclusionTable = paste0(cohortTable, "_inclusion"),
  cohortInclusionResultTable = paste0(cohortTable, "_inclusion_result"),
  cohortInclusionStatsTable = paste0(cohortTable, "_inclusion_stats"),
  cohortSummaryStatsTable = paste0(cohortTable, "_summary_stats")
)
```

Arguments

`connectionDetails`
An object of type `connectionDetails` as created using the `createConnectionDetails` function in the DatabaseConnector package. Can be left NULL if `connection` is provided.

`connection`
An object of type `connection` as created using the `connect` function in the DatabaseConnector package. Can be left NULL if `connectionDetails` is provided, in which case a new connection will be opened at the start of the function, and closed when the function finishes.

`resultsDatabaseSchema`
Schema name where the statistics tables reside. Note that for SQL Server, this should include both the database and schema name, for example 'scratch.dbo'.

`cohortId`
The cohort definition ID used to reference the cohort in the cohort table.

`simplify`
Simply output the attrition table?

`cohortTable`
Name of the cohort table. Used only to conveniently derive names of the four rule statistics tables.

`cohortInclusionTable`
Name of the inclusion table, one of the tables for storing inclusion rule statistics.

`cohortInclusionResultTable`
Name of the inclusion result table, one of the tables for storing inclusion rule statistics.

`cohortInclusionStatsTable`
Name of the inclusion stats table, one of the tables for storing inclusion rule statistics.

`cohortSummaryStatsTable`
Name of the summary stats table, one of the tables for storing inclusion rule statistics.

Value

If `simplify = TRUE`, this function returns a single data frame. Else a list of data frames is returned.

getInclusionStatisticsFromFiles

Get inclusion criteria statistics from files

Description

Gets inclusion criteria statistics from files, as stored when using the ROhdsiWebApi::insertCohortDefinitionSetInPa function with generateStats = TRUE.

Usage

```
getInclusionStatisticsFromFiles(
  cohortId,
  folder,
  cohortInclusionFile = file.path(folder, "cohortInclusion.csv"),
  cohortInclusionResultFile = file.path(folder, "cohortIncResult.csv"),
  cohortInclusionStatsFile = file.path(folder, "cohortIncStats.csv"),
  cohortSummaryStatsFile = file.path(folder, "cohortSummaryStats.csv"),
  simplify = TRUE
)
```

Arguments

cohortId The cohort definition ID used to reference the cohort in the cohort table.

folder The path to the folder where the inclusion statistics are stored.

cohortInclusionFile Name of the inclusion table, one of the tables for storing inclusion rule statistics.

cohortInclusionResultFile Name of the inclusion result table, one of the tables for storing inclusion rule statistics.

cohortInclusionStatsFile Name of the inclusion stats table, one of the tables for storing inclusion rule statistics.

cohortSummaryStatsFile Name of the summary stats table, one of the tables for storing inclusion rule statistics.

simplify Simply output the attrition table?

Value

If simplify = TRUE, this function returns a single data frame. Else a list of data frames is returned.

instantiateCohortSet *Instantiate a set of cohort*

Description

This function instantiates a set of cohort in the cohort table, using definitions that are fetched from a WebApi interface. Optionally, the inclusion rule statistics are computed and stored in the inclusionStatisticsFolder.

Usage

```
instantiateCohortSet(
  connectionDetails = NULL,
  connection = NULL,
  cdmDatabaseSchema,
  oracleTempSchema = NULL,
  cohortDatabaseSchema = cdmDatabaseSchema,
  cohortTable = "cohort",
  cohortIds = NULL,
  minCellCount,
  generateInclusionStats = FALSE,
  inclusionStatisticsFolder = NULL,
  createCohortTable = FALSE,
  incremental = FALSE,
  incrementalFolder = NULL
)
```

Arguments

connectionDetails
: An object of type connectionDetails as created using the createConnectionDetails function in the DatabaseConnector package. Can be left NULL if connection is provided.

connection
: An object of type connection as created using the connect function in the DatabaseConnector package. Can be left NULL if connectionDetails is provided, in which case a new connection will be opened at the start of the function, and closed when the function finishes.

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 privileges for storing temporary tables.

cohortDatabaseSchema
: Schema name where your cohort table resides. Note that for SQL Server, this should include both the database and schema name, for example 'scratch.dbo'.

cohortTable
: Name of the cohort table.

cohortIds
: Optionally, provide a subset of cohort IDs to restrict the construction to.

`generateInclusionStats`
Compute and store inclusion rule statistics?

`inclusionStatisticsFolder`
The folder where the inclusion rule statistics are stored. Can be left NULL if `generateInclusionStats = FALSE`.

`createCohortTable`
Create the cohort table? If `incremental = TRUE` and the table already exists this will be skipped.

`incremental` Create only cohorts that haven't been created before?

`incrementalFolder`
If `incremental = TRUE`, specify a folder where records are kept of which definition has been executed.

`packageName` The name of the package containing the cohort definitions. Can be left NULL if `baseUrl` and `cohortSetReference` have been specified.

`cohortToCreateFile`
The location of the cohortToCreate file within the package. Is ignored if `baseUrl` and `cohortSetReference` have been specified.

`baseUrl` The base URL for the WebApi instance, for example: "http://server.org:80/WebAPI". Can be left NULL if `packageName` and `cohortToCreateFile` have been specified.

`cohortSetReference`
A data frame with four columns, as described in the details. Can be left NULL if `packageName` and `cohortToCreateFile` have been specified.

Details

Currently two ways of executing this function are supported, either (1) embedded in a study package, assuming the cohort definitions are stored in that package using the `ROhdsiWebApi::insertCohortDefinitionSet` or (2) by using a WebApi interface to retrieve the cohort definitions.

When using this function from within a study package, use the `packageName` and `cohortToCreateFile` to specify the name of the study package, and the name of the cohortToCreate file within that package, respectively

When using this function using a WebApi interface, use the `baseUrl` and `cohortSetReference` to specify how to connect to the WebApi, and which cohorts to fetch, respectively.

The `cohortSetReference` argument must be a data frame with the following columns:

atlasId The cohort ID in ATLAS.

atlasName The full name of the cohort. This will be shown in the Shiny app.

cohortId The cohort ID to use in the package. Usually the same as the cohort ID in ATLAS.

name A short name for the cohort, to use to create file names. do not use special characters.

preMergeDiagnosticsFiles *Premerge Shiny diagnostics files*

Description

If there are many diagnostics files, starting the Shiny app may take a very long time. This function already does most of the preprocessing, increasing loading speed.

The merged data will be stored in the same folder, and will automatically be recognized by the Shiny app.

Usage

```
preMergeDiagnosticsFiles(dataFolder)
```

Arguments

`dataFolder` folder where the exported zip files for the diagnostics are stored. Use the `runCohortDiagnostics` function to generate these zip files. Zip files containing results from multiple databases can be placed in the same folder.

preMergeResultsFiles *Premerge Shiny results files*

Description

If there are many results files, starting the Shiny app may take a very long time. This function already does most of the preprocessing, increasing loading speed.

The merged data will be stored in the same folder, and will automatically be recognized by the Shiny app.

Usage

```
preMergeResultsFiles(dataFolder)
```

Arguments

`dataFolder` folder where the exported zip files for the diagnostics are stored. Use the runStudy function to generate these zip files. Zip files containing results from multiple databases can be placed in the same folder.

`uploadResults` *Upload results to OHDSI server*

Description

Upload results to OHDSI server

Usage

```
uploadResults(outputFolder, privateKeyFileName, userName, remoteFolder)
```

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.

`privateKeyFileName` A character string denoting the path to the RSA private key provided by the study coordinator.

`userName` A character string containing the user name provided by the study coordinator.

Details

This function uploads 'Results_\<databaseId\>.zip' to the OHDSI SFTP server. Before sending, you can inspect the zip file, which contains (zipped) CSV files. You can send the zip file from a different computer than the one on which is was created.

`writeBulkCharacteristics` *Write cohort characteristics in bulk to the file system*

Description

This function will retrieve the results from the temp tables created in createBulkCharacteristics

Usage

```
writeBulkCharacteristics(
  connection,
  oracleTempSchema,
  counts,
  minCellCount,
  databaseId,
  exportFolder
)
```

Index