How to Use PhenotypeLibrary R Package

Gowtham A. Rao

2025-03-29

Contents

PhenotypeLibrary is part of HADES

1 Installation

- This is an installable R-package that may be installed as follows:

```
remotes::install_github("OHDSI/PhenotypeLibrary")
```

2 Retrieval

- The list of cohort definitions available may be retrieved as follows:

```
PhenotypeLibrary::getPhenotypeLog()
```

```
#> # A tibble: 661 x 88
#>    cohortId cohortName           cohortNameAtlas cohortNameFormatted cohortNameLong librarian status
#>       <dbl> <chr>                <chr>           <chr>               <chr>          <chr>     <chr>
#>  1        3 Cough or Sputum      [P] Cough or S~ Cough or Sputum     Cough or Sput~ rao@ohds~ Pendi~
#>  2        4 Diarrhea             [P] Diarrhea    Diarrhea            Diarrhea       rao@ohds~ Pendi~
#>  3        5 Dyspnea              [P] Dyspnea     Dyspnea             Dyspnea        rao@ohds~ Pendi~
#>  4        6 Fever                [P] Fever       Fever               Fever          rao@ohds~ Pendi~
#>  5        7 Headache or Headach~ [P] Headache o~ Headache or Headac~ Headache or H~ rao@ohds~ Pendi~
#>  6        8 Altered smell or ta~ [P] Altered sm~ Altered smell or t~ Altered smell~ rao@ohds~ Pendi~
#>  7        9 Sore throat          [P] Sore throat Sore throat         Sore throat    rao@ohds~ Pendi~
#>  8       10 Nausea or Vomiting   [P] Nausea or ~ Nausea or Vomiting  Nausea or Vom~ rao@ohds~ Pendi~
#>  9       11 Malaise and or fati~ [P] Malaise an~ Malaise and or fat~ Malaise and o~ rao@ohds~ Pendi~
#> 10       12 Rhinitis or common ~ [P] Rhinitis o~ Rhinitis or common~ Rhinitis or c~ rao@ohds~ Pendi~
#> # i 651 more rows
```

```
#> # i 78 more variables: isCirceJson <dbl>, contributors <chr>, contributorOrcIds <chr>, contributorOr
#> #   peerReviewers <chr>, peerReviewerOrcIds <dbl>, recommendedReferentConceptIds <chr>, ohdsiForumPo
#> #   modifiedDate <date>, lastModifiedBy <dbl>, replaces <dbl>, notes <chr>, isReferenceCohort <dbl>,
#> #   censorWindowEndDate <dbl>, collapseSettingsType <chr>, collapseEraPad <dbl>, exitStrategy <chr>,
#> #   exitDateOffSet <dbl>, numberOfInclusionRules <dbl>, initialEventLimit <chr>, initialEventRestric
#> #   initialEventRestrictionAdditionalCriteriaLimit <chr>, inclusionRuleQualifyingEventLimit <chr>, nu
```

- You can extract one or more cohort definitions into a cohortDefinitionSet object as

```
cohortDefinitionSet <- PhenotypeLibrary::getPlCohortDefinitionSet(cohortIds = c(1, 2, 3))

cohortDefinitionSet
```

```
#> # A tibble: 2 x 4
#>   cohortId cohortName                                    json
#>      <dbl> <chr>                                       <chr>
#> 1        2 COVID-19 diagnosis or SARS-CoV-2 test (1pos) "{\n\t\"cdmVersionRange\" : \">=5.0.0\",\n\t
#> 2        3 Cough or Sputum                              "{\n\t\"cdmVersionRange\" : \">=5.0.0\",\n\t
```

- cohortDefinitionSet is now a data.frame with specifications for the cohort ids 1, 2 and 3. For cohorts that conform to OHDSI Circe specifications, the field json is the cohort json specification that may be posted into your Atlas instance. The SQL is the SQL rendered from the JSON. For cohorts that do not conform to OHDSI Circe specification, only the SQL is provided and the json is left empty.

3 Use

- You can instantiate the cohorts in your environment as follows using (OHDSI/CohortGenerator)[https://github.com/OHDSI/CohortGenerator].

```
connectionDetails <-
  DatabaseConnector::createConnectionDetails(
    dbms = "postgresql",
    server = "some.server.com/ohdsi",
    user = "joe",
    password = "secret"
  )
cdmDatabaseSchema <- "cdm_synpuf"
cohortDatabaseSchema <- "scratch.dbo"
cohortTables <- CohortGenerator::getCohortTableNames()
CohortGenerator::generateCohortSet(
  connectionDetails = connectionDetails,
  cdmDatabaseSchema = cdmDatabaseSchema,
  cohortDatabaseSchema = cohortDatabaseSchema,
  cohortTableNames = cohortTables,
  cohortDefinitionSet = cohortDefinitionSet
)
```

- You can also run cohort diagnostics on this cohortDefinitionSet object as follows:

```r
databaseId <- "synpuf"

databaseName <-
  "Medicare Claims Synthetic Public Use Files (SynPUFs)"

databaseDescription <-
  "Medicare Claims Synthetic Public Use Files (SynPUFs) were created to allow interested parties to gai

CohortDiagnostics::executeDiagnostics(
  cohortDefinitionSet = cohortDefinitionSet,
  exportFolder = outputFolder,
  databaseId = databaseId,
  databaseName = databaseName,
  databaseDescription = databaseDescription,
  cohortDatabaseSchema = cohortDatabaseSchema,
  cdmDatabaseSchema = cdmDatabaseSchema,
  connectionDetails = connectionDetails,
  cohortTableNames = cohortTableNames
)
```