Using the cohort covariate code

Jenna M. Reps

2020-10-13

Contents

1 Introduction

This vignette describes how one can use the function 'createCohortCovariateSettings' to define covariates using the cohort table. The first steps are to create all required phenotypes into a cohort table and make a note of the cohort database schema, cohort table and cohort definition ids.

1.1 createCohortCovariateSettings

This function contains the settings required to define the covariate. For a cohort covariate, the code will check the cohort table 'cohortDatabaseSchema'.'cohortTable' to find all rows where the column cohort_definition_id is 'cohortId'. It will then check whether the cohort_start_date column calls between the index date plus the 'startDay' and the index date plus the 'endDay'. If 'count' is set to TRUE then it will return the number of rows where the cohort_start_date falls between the start and end dates, otherwise it will return 1 if there are rows and 0 else. The settings 'ageInteraction' and 'lnAgeInteraction' enable the user to create age/ln(age) interaction terms. Finally, the 'analysisId' is used to create the cohort covariateId as 1000*'cohortId' + 'analysisId'.

Table 1: The inputs into the create function

Input	Description
covariateName	The name of the covariate
covariateId	The id of the covariate - generally cohortId*1000+analysisId
cohortDatabaseSchema	The database schema with the cohort used to create a covariate
cohortTable	The table with the cohort used to create a covariate
cohortId	The cohort definition id for the cohort used to create a covariate
startDay	How many days prior to index to see whether the covariate cohort occurs after
endDay	How many days relative to index to see whether the covariate cohort occurs before
count	Count how many unique dates occur in the covariate cohort during the start and end dates for ea
ageInteraction	Include interaction with age
lnAgeInteraction	Include interaction with ln(age)
analysisId	The analysis id for the covariate

1.2 Example

Assuming a cohort table exists at: 'the_cohortDatabaseSchema'.'cohort' and contains the set of dates a patient is initially diagnosed with diabetes with the cohort_definition_id of 999. To create a diabetes covariate using a cohort definition run:

```
cohortCov1 <- createCohortCovariateSettings(covariateName = 'example diabetes anytime prior',
                                            analysisId = 456,
                                            covariateId = 999*1000+456,
                                        cohortDatabaseSchema = 'the_cohortDatabaseSchema'
                                        cohortTable = 'cohort',
                                        cohortId = 999,
                                        startDay= -9999,
                                        endDay=-1,
                                        count= FALSE),
                                        ageInteraction = FALSE,
                                        lnAgeInteraction = FALSE)
```

If you wanted to only find patients diagnosed with diabetes in the past 365 days you can use:

```
cohortCov2 <- createCohortCovariateSettings(covariateName = 'example diabetes within 365 days prior',
                                            analysisId = 456,
                                            covariateId = 999*1000+456,
                                        cohortDatabaseSchema = 'the_cohortDatabaseSchema'
                                        cohortTable = 'cohort',
                                        cohortId = 999,
                                        startDay= -365,
                                        endDay=-1,
                                        count= FALSE),
                                        ageInteraction = FALSE,
                                        lnAgeInteraction = FALSE)
```

To include an age interaction (age in years or 0 is the value rather than 1 or 0):

```
cohortCov3 <- createCohortCovariateSettings(covariateName = 'example diabetes anytime prior interaction
                                            analysisId = 456,
                                            covariateId = 999*1000+456,
                                        cohortDatabaseSchema = 'the_cohortDatabaseSchema'
                                        cohortTable = 'cohort',
                                        cohortId = 999,
                                        startDay= -9999,
                                        endDay=-1,
                                        count= FALSE),
                                        ageInteraction = TRUE,
                                        lnAgeInteraction = FALSE)
```

To include an ln(age) interaction (natural log of age in years or 0 is the value rather than 1 or 0):

```
cohortCov4 <- createCohortCovariateSettings(covariateName = 'example diabetes anytime prior interaction
                                            analysisId = 456,
                                            covariateId = 999*1000+456,
                                        cohortDatabaseSchema = 'the_cohortDatabaseSchema'
                                        cohortTable = 'cohort',
                                        cohortId = 999,
                                        startDay= -9999,
                                        endDay=-1,
                                        count= FALSE),
```

```
                                    ageInteraction = FALSE,
                                    lnAgeInteraction = TRUE)
```

To include all the above as covariates, combine them into a list:

```
cohortCov <- list(cohortCov1,cohortCov2,cohortCov3,cohortCov4)
```