Single studies using the CohortMethod package

Martijn J. Schuemie, Marc A. Suchard and Patrick Ryan

2024-05-31

Contents

1	Introduction	1
2	Installation instructions	2
3	Data extraction 3.1 Configuring the connection to the server 3.2 Preparing the exposures and outcome(s) 3.3 Extracting the data from the server	2 2 2 4
4	Defining the study population	6
5	Propensity scores5.1Fitting a propensity model5.2Propensity score diagnostics5.3Using the propensity score5.4Evaluating covariate balance5.5Inspecting select population characteristics5.6Generalizability	7 7 9 12 14 15
6	Follow-up and power	15
7	Outcome models 7.1 Fitting a simple outcome model	16 17 18 19 19 19 20
8	Acknowledgments	21

1 Introduction

This vignette describes how you can use the CohortMethod package to perform a single new-user cohort study. We will walk through all the steps needed to perform an exemplar study, and we have selected the well-studied topic of the effect of coxibs versus non-selective non-steroidal anti-inflammatory drugs (NSAIDs) on gastrointestinal (GI) bleeding-related hospitalization. For simplicity, we focus on one coxib – celecoxib – and one non-selective NSAID – diclofenac.

2 Installation instructions

Before installing the CohortMethod package make sure you have Java available. For Windows users, RTools is also necessary. See these instructions for properly configuring your R environment.

The CohortMethod package is currently maintained in a Github repository, and has dependencies on other packages in Github. All of these packages can be downloaded and installed from within R using the drat package:

```
install.packages("remotes")
remotes::install_github("ohdsi/CohortMethod")
```

Once installed, you can type library(CohortMethod) to load the package.

3 Data extraction

The first step in running the CohortMethod is extracting all necessary data from the database server holding the data in the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) format.

3.1 Configuring the connection to the server

We need to tell R how to connect to the server where the data are. CohortMethod uses the DatabaseConnector package, which provides the createConnectionDetails function. Type ?createConnectionDetails for the specific settings required for the various database management systems (DBMS). For example, one might connect to a PostgreSQL database using this code:

The last few lines define the cdmDatabaseSchema, cohortDatabaseSchema, and cohortTable variables. We'll use these later to tell R where the data in CDM format live, and where we want to write intermediate tables. Note that for Microsoft SQL Server, databaseschemas need to specify both the database and the schema, so for example cdmDatabaseSchema <- "my_cdm_data.dbo". For database platforms that do not support temp tables, such as Oracle, it is also necessary to provide a schema where the user has write access that can be used to emulate temp tables. PostgreSQL supports temp tables, so we can set options(sqlRenderTempEmulationSchema = NULL) (or not set the sqlRenderTempEmulationSchema at all.)

3.2 Preparing the exposures and outcome(s)

We need to define the exposures and outcomes for our study. Here, we will define our exposures using the OHDSI Capr package. We define two cohorts, one for celecoxib and one for diclofenac. For each cohort we require a prior diagnosis of 'osteoarthritis of knee', and 365 days of continuous prior observation. we restrict to the first exposure per person:

```
library(Capr)
```

```
osteoArthritisOfKneeConceptId <- 4079750
celecoxibConceptId <- 1118084
```

```
diclofenacConceptId <- 1124300
osteoArthritisOfKnee <- cs(</pre>
  descendants(osteoArthritisOfKneeConceptId),
  name = "Osteoarthritis of knee"
)
attrition = attrition(
  "prior osteoarthritis of knee" = withAll(
    atLeast(1, conditionOccurrence(osteoArthritisOfKnee), duringInterval(eventStarts(-Inf, 0)))
  )
)
celecoxib <- cs(</pre>
  descendants(celecoxibConceptId),
 name = "Celecoxib"
)
diclofenac <- cs(
  descendants(diclofenacConceptId),
  name = "Diclofenac"
)
celecoxibCohort <- cohort(</pre>
  entry = entry(
    drugExposure(celecoxib, firstOccurrence()),
    observationWindow = continuousObservation(priorDays = 365)
 ),
 attrition = attrition,
  exit = exit(endStrategy = drugExit(celecoxib,
                                      persistenceWindow = 30,
                                      surveillanceWindow = 0))
)
diclofenacCohort <- cohort(</pre>
  entry = entry(
    drugExposure(diclofenac, firstOccurrence()),
    observationWindow = continuousObservation(priorDays = 365)
  ).
  attrition = attrition,
  exit = exit(endStrategy = drugExit(diclofenac,
                                      persistenceWindow = 30,
                                      surveillanceWindow = 0))
)
# Note: this will automatically assign cohort IDs 1 and 2, respectively:
exposureCohorts <- makeCohortSet(celecoxibCohort, diclofenacCohort)</pre>
```

We'll pull the outcome definition from the OHDSI PhenotypeLibrary:

```
library(PhenotypeLibrary)
outcomeCohorts <- getPlCohortDefinitionSet(77) # GI bleed</pre>
```

We combine the exposure and outcome cohort definitions, and use CohortGenerator to generate the cohorts:

If all went well, we now have a table with the cohorts of interest. We can see how many entries per cohort:

```
connection <- DatabaseConnector::connect(connectionDetails)
sql <- "SELECT cohort_definition_id, COUNT(*) AS count FROM @cohortDatabaseSchema.@cohortTable GROUP BY
DatabaseConnector::renderTranslateQuerySql(connection, sql, cohortDatabaseSchema = cohortDatabaseSchem
DatabaseConnector::disconnect(connection)</pre>
```

```
## cohort_concept_id count
## 1 1 109307
## 2 2 176675
## 3 77 733601
```

3.3 Extracting the data from the server

Now we can tell CohortMethod to extract the cohorts, construct covariates, and extract all necessary data for our analysis.

Important: The target and comparator drug must not be included in the covariates, including any descendant concepts. You will need to manually add the drugs and descendants to the excludedCovariateConceptIds of the covariate settings. In this example code we exclude the concepts for celecoxib and diclofenac and specify addDescendantsToExclude = TRUE:

```
# Define which types of covariates must be constructed:
covSettings <- createDefaultCovariateSettings(</pre>
  excludedCovariateConceptIds = c(diclofenacConceptId, celecoxibConceptId),
  addDescendantsToExclude = TRUE
)
#Load data:
cohortMethodData <- getDbCohortMethodData(</pre>
  connectionDetails = connectionDetails,
  cdmDatabaseSchema = cdmDatabaseSchema,
  targetId = 1,
  comparatorId = 2,
  outcomeIds = 77,
  exposureDatabaseSchema = cohortDatabaseSchema,
  exposureTable = cohortTable,
  outcomeDatabaseSchema = cohortDatabaseSchema,
  outcomeTable = cohortTable,
  covariateSettings = covSettings
)
cohortMethodData
## # CohortMethodData object
##
## Target cohort ID: 1
## Comparator cohort ID: 2
## Outcome cohort ID(s): 77
##
## Inherits from CovariateData:
## # CovariateData object
```

```
##
##
##
##
##
##
## Inherits from Andromeda:
## # Andromeda object
## # Andromeda object
## # Physical location: C:\Users\admin_mschuemi\AppData\Local\Temp\2\RtmpWSqLkf\file2c305f401b93.sqlite
##
## Tables:
## $analysisRef (analysisId, analysisName, domainId, startDay, endDay, isBinary, missingMeansZero)
## $cohorts (rowId, personSeqId, personId, treatment, cohortStartDate, daysFromObsStart, daysToCohortEnd, daysToObsEnd)
## $covariateRef (covariateId, covariateName, analysisId, conceptId)
## $covariates (rowId, covariateId, covariateValue)
## $outcomes (rowId, outcomeId, daysToEvent)
```

There are many parameters, but they are all documented in the CohortMethod manual. The createDefaultCovariateSettings function is described in the FeatureExtraction package. In short, we are pointing the function to the table created earlier and indicating which concept IDs in that table identify the target, comparator and outcome. We instruct that the default set of covariates should be constructed, including covariates for all conditions, drug exposures, and procedures that were found on or before the index date. To customize the set of covariates, please refer to the FeatureExtraction package vignette by typing vignette("UsingFeatureExtraction", package="FeatureExtraction").

All data about the cohorts, outcomes, and covariates are extracted from the server and stored in the cohortMethodData object. This object uses the Andromeda package to store information in a way that ensures R does not run out of memory, even when the data are large. We can use the generic summary() function to view some more information of the data we extracted:

summary(cohortMethodData)

```
## CohortMethodData object summary
##
## Target cohort ID: 1
## Comparator cohort ID: 2
## Outcome cohort ID(s): 77
##
## Target persons: 109307
## Comparator persons: 176675
##
## Outcome counts:
##
      Event count Person count
## 77
            37779
                         27913
##
## Covariates:
## Number of covariates: 78589
## Number of non-zero covariate values: 142730640
```

3.3.1 Saving the data to file

Creating the cohortMethodData file can take considerable computing time, and it is probably a good idea to save it for future sessions. Because cohortMethodData uses Andromeda, we cannot use R's regular save function. Instead, we'll have to use the saveCohortMethodData() function:

saveCohortMethodData(cohortMethodData, "coxibVsNonselVsGiBleed.zip")

We can use the loadCohortMethodData() function to load the data in a future session.

3.3.2 Defining new users

Typically, a new user is defined as first time use of a drug (either target or comparator), and typically a washout period (a minimum number of days prior first use) is used to make sure it is truly first use. When

using the CohortMethod package, you can enforce the necessary requirements for new use in three ways:

- 1. When creating the cohorts in the database, for example using Capr.
- 2. When loading the cohorts using the getDbCohortMethodData function, you can use the firstExposureOnly, removeDuplicateSubjects, restrictToCommonPeriod, and washoutPeriod arguments. (As shown in the example above).
- 3. When defining the study population using the createStudyPopulation function (see below) using the firstExposureOnly, removeDuplicateSubjects, restrictToCommonPeriod, and washoutPeriod arguments.

The advantage of option 1 is that the input cohorts are already fully defined outside of the CohortMethod package, and for example external cohort characterization tools can be used on the same cohorts used in this package. The advantage of options 2 and 3 is that it saves you the trouble of limiting to first use yourself, for example allowing you to directly use the drug_era table in the CDM. Option 2 is more efficient than 3, since only data for first use will be fetched, while option 3 is less efficient but allows you to compare the original cohorts to the study population.

4 Defining the study population

Typically, the exposure cohorts and outcome cohorts will be defined independently of each other. When we want to produce an effect size estimate, we need to further restrict these cohorts and put them together, for example by removing exposed subjects that had the outcome prior to exposure, and only keeping outcomes that fall within a defined risk window. For this we can use the createStudyPopulation function:

```
studyPop <- createStudyPopulation(
    cohortMethodData = cohortMethodData,
    outcomeId = 3,
    firstExposureOnly = FALSE,
    restrictToCommonPeriod = FALSE,
    washoutPeriod = 0,
    removeDuplicateSubjects = "keep all",
    removeSubjectsWithPriorOutcome = TRUE,
    minDaysAtRisk = 1,
    riskWindowStart = 0,
    startAnchor = "cohort start",
    riskWindowEnd = 30,
    endAnchor = "cohort end"
)
```

Note that we've set firstExposureOnly and removeDuplicateSubjects to FALSE, and washoutPeriod to zero because we already filtered on these arguments when using the getDbCohortMethodData function. During loading we set restrictToCommonPeriod to FALSE, and we do the same here because we do not want to force the comparison to restrict only to time when both drugs are recorded. We specify the outcome ID we will use, and that people with outcomes prior to the risk window start date will be removed. The risk window is defined as starting at the cohort start date (the index date, riskWindowStart = 0 and startAnchor = "cohort start"), and the risk windows ends 30 days after the cohort ends (riskWindowEnd = 30 and endAnchor = "cohort end"). Note that the risk windows are truncated at the end of observation or the study end date. We also remove subjects who have no time at risk. To see how many people are left in the study population we can always use the getAttritionTable function:

getAttritionTable(studyPop)

##	#	A tibble: 4 x 5				
##		description	targetPersons	${\tt comparator} {\tt Persons}$	targetExposures	comparatorExposures
##		<chr></chr>	<int></int>	<int></int>	<int></int>	<int></int>
##	1	Original cohorts	109307	176675	109307	176675
##	2	Restricting duplicate subje	99734	164277	99734	164277

## 3 No prior outcome	96089	156060	96089	156060
## 4 Have at least 1 days at ris	96050	155990	96050	155990

One additional filtering step that is often used is matching or trimming on propensity scores, as will be discussed next.

5 Propensity scores

The CohortMethod can use propensity scores to adjust for potential confounders. Instead of the traditional approach of using a handful of predefined covariates, CohortMethod typically uses thousands to millions of covariates that are automatically constructed based on conditions, procedures and drugs in the records of the subjects.

5.1 Fitting a propensity model

We can fit a propensity model using the covariates constructed by the getDbcohortMethodData() function:

```
ps <- createPs(cohortMethodData = cohortMethodData, population = studyPop)</pre>
```

The createPs() function uses the Cyclops package to fit a large-scale regularized logistic regression.

To fit the propensity model, Cyclops needs to know the hyperparameter value which specifies the variance of the prior. By default Cyclops will use cross-validation to estimate the optimal hyperparameter. However, be aware that this can take a really long time. You can use the prior and control parameters of the createPs() to specify Cyclops behavior, including using multiple CPUs to speed-up the cross-validation.

5.2 Propensity score diagnostics

We can compute the area under the receiver-operator curve (AUC) for the propensity score model:

computePsAuc(ps)

[1] 0.81

We can also plot the propensity score distribution, although we prefer the preference score distribution:

plotPs(ps,

```
scale = "preference",
showCountsLabel = TRUE,
showAucLabel = TRUE,
showEquiposeLabel = TRUE)
```



It is also possible to inspect the propensity model itself by showing the covariates that have non-zero coefficients:

getPsModel(ps, cohortMethodData)

```
## # A tibble: 6 x 3
##
     coefficient covariateId covariateName
##
           <dbl>
                        <dbl> <chr>
## 1
           -3.49
                  1150871413 ... gh 0 days relative to index: misoprostol
## 2
            3.27
                      2001006 index year: 2001
                      2002006 index year: 2002
## 3
            3.09
            2.70
                      2003006 index year: 2003
## 4
                      2004006 index year: 2004
## 5
            2.51
## 6
            1.79
                      2007006 index year: 2007
```

One advantage of using the regularization when fitting the propensity model is that most coefficients will shrink to zero and fall out of the model. It is a good idea to inspect the remaining variables for anything that should not be there, for example variations of the drugs of interest that we forgot to exclude.

Finally, we can inspect the percent of the population in equipoise, meaning they have a preferce score between 0.3 and 0.7:

CohortMethod::computeEquipoise(ps)

[1] 0.4900214

A low equipoise indicates there is little overlap between the target and comparator populations.

5.3 Using the propensity score

We can use the propensity scores to trim, stratify, match, or weigh our population. For example, one could trim to equipoise, meaning only subjects with a preference score between 0.25 and 0.75 are kept:

```
trimmedPop <- trimByPsToEquipoise(ps)
plotPs(trimmedPop, ps, scale = "preference")</pre>
```

```
## Population size after trimming is 123505
```



Instead (or additionally), we could stratify the population based on the propensity score: stratifiedPop <- stratifyByPs(ps, numberOfStrata = 5) plotPs(stratifiedPop, ps, scale = "preference")



We can also match subjects based on propensity scores. In this example, we're using one-to-one matching: matchedPop <- matchOnPs(ps, caliper = 0.2, caliperScale = "standardized logit", maxRatio = 1)
plotPs(matchedPop, ps)</pre>

Population size after matching is 115832



Note that for both stratification and matching it is possible to specify additional matching criteria such as age and sex using the stratifyByPsAndCovariates() and matchOnPsAndCovariates() functions, respectively.

We can see the effect of trimming and/or matching on the population using the getAttritionTable function:

```
getAttritionTable(matchedPop)
```

##	#	A tibble: 5 x 5				
##		description	targetPersons	comparator Persons	targetExposures	comparatorExposures
##		<chr></chr>	<int></int>	<int></int>	<int></int>	<int></int>
##	1	Original cohorts	109307	176675	109307	176675
##	2	Restricting duplicate subje	99734	164277	99734	164277
##	3	No prior outcome	96089	156060	96089	156060
##	4	Have at least 1 days at ris	96050	155990	96050	155990
##	5	Matched on propensity score	57916	57916	57916	57916

Or, if we like, we can plot an attrition diagram:

drawAttritionDiagram(matchedPop)



5.4 Evaluating covariate balance

To evaluate whether our use of the propensity score is indeed making the two cohorts more comparable, we can compute the covariate balance before and after trimming, matching, and/or stratifying:

balance <- computeCovariateBalance(matchedPop, cohortMethodData)</pre>

plotCovariateBalanceScatterPlot(balance, showCovariateCountLabel = TRUE, showMaxLabel = TRUE)

Warning: Removed 27794 rows containing missing values (`geom_point()`).







The 'before matching' population is the population as extracted by the getDbCohortMethodData function,

so before any further filtering steps.

5.5 Inspecting select population characteristics

It is customary to include a table in your paper that lists some select population characteristics before and after matching/stratification/trimming. This is usually the first table, and so will be referred to as 'table 1'. To generate this table, you can use the createCmTable1 function:

createCmTable1(balance)

Target Comparator Target Comparator Agricular X Bid. 4117 X Bid. 4117 X Agricular 0.0 0.0 0.0 0.0 0.0 0.0 0.0 B - 49 0.0 <		Before matching			After matching		
Lag Cont Cont <thc< td=""><td>(have at weight in</td><td>Target</td><td>Comparator</td><td>0+3 3:66</td><td>Target</td><td>Comparator</td><td>0+3 3:66</td></thc<>	(have at weight in	Target	Comparator	0+3 3:66	Target	Comparator	0+3 3:66
0 0	Age group	7.	7.	Sta. diff	7.	7.	Std. diff
45 49 0.0 0.0 0.00 0.00 0.00 50 50 60 0.00 0.00 0.00 0.00 60 64 1.0 1.3 0.02 1.2 1.2 0.00 61 67 77 2.2 1.2 0.00 2.5 2.2.3 0.00 60 74 0.00 2.1.4 0.00 2.1.4 0.00 0.00 0.00 0.00 60 74 0.00 2.1.4 0.00 2.1.4 0.00	40 - 44	0.0	0.0	0.00	0.0	0.0	0.00
bb bb<	45 - 49	0.0	0.0	0.00	0.0	0.0	0.00
00 62 90 1.2 1.2 1.2 1.2 1.2 1.2 1.2 1.2 1.2 1.2 1.2 1.2 1.2 1.2 0.00 76 79 2.2 1.2 0.00 2.1 1.2 0.00 1.2 1.2 0.00 80 84 1.5 1.6 0.00 1.5 1.5 0.00 85 84 7.7 9.3 0.00 0.6 0.6 0.00 95 95 95 0.00 0.0 <	50 - 54 55 - 59	0.2	0.2	-0.01	0.2	0.2	-0.01
65 69 22.2 19.2 0.07 22.5 22.3 0.00 70 74 21.5 0.08 27.4 21.5 0.00 80 93 0.51 15.1 16.2 0.00 15.1 16.1 0.00 80 93 0.5 1.6 1.6 0.00 0.1 0.1 100 104 0.1 0.1 0.1 0.0 0.0 0.0 0.0 100 104 0.0 0.0 0.0 0.0 0.0 0.0 105 105 0.0 0.0 0.0 0.0 0.0 0.0 105 105 0.00 0.0 0.0 0.0 0.0 0.0 105 105 0.00 0.1 1.4 0.00 0.1 0.00 106 11.4 0.00 11.4 0.00 0.01 0.00 105 0.00 1.1 0.00 1.1 0.00 0.0	60 - 64	1.0	1.3	-0.02	1.2	1.2	0.00
70 74 28.2 25.9 0.06 27.4 27.2 0.00 80 94 25.1 27.2 0.03 15.2 17.3 0.00 80 94 26.6 4.0 -0.08 8.1 8.1 0.00 80 -94 2.6 4.0 -0.08 8.1 8.1 0.00 105 -103 0.0 0.0 0.0 0.0 0.0 0.0 0.0 105 -104 0.1 1.0 -0.22 0.0 0.0 0.0 0.0 105 -105 0.00 0.2 0.2 0.0 0.0 0.0 0.0 Chardi Liver disease 2.7 2.6 -0.07 2.1 2.0 0.	65 - 69	22.2	19.2	0.07	22.5	22.3	0.00
ho -9 22.1 21.6 0.13 21.4 21.6 0.00 S5 -80 -7.7 5.9 -0.08 2.9 2.9 0.00 90 -94 2.6 4.0 -0.08 2.9 2.9 0.00 90 -94 0.13 1.1 0.10 0.06 0.0 0.00 0.0 0.00 <td>70 - 74</td> <td>28.2</td> <td>25.9</td> <td>0.05</td> <td>27.4</td> <td>27.2</td> <td>0.00</td>	70 - 74	28.2	25.9	0.05	27.4	27.2	0.00
bs bs<	75 - 79	22.1	21.6	0.01	21.4	21.6	0.00
00 - 54 2.6 4.0 -0.08 2.9 2.0 0.00 105 - 99 0.6 0.0 </td <td>85 - 89</td> <td>7.7</td> <td>9.9</td> <td>-0.03</td> <td>8.1</td> <td>8.1</td> <td>0.00</td>	85 - 89	7.7	9.9	-0.03	8.1	8.1	0.00
98 0.5 1.0 -0.68 0.6 0.6 0.0 0.0 100 130	90 - 94	2.6	4.0	-0.08	2.9	2.9	0.00
100 101 -0.02 0.0 0.1 -0.01 Redical history: General 64.2 67.3 -0.02 66.2 66.1 0.00 Medical history: General 2.7 25.6 -0.07 23.9 23.8 0.00 Acture respiratory disease 0.1 0.2 0.00 1.1 0.2 0.00 Chronic obseque 0.4 0.6 0.6 1.5 0.6 1.5 0.6 1.5 0.6 1.5 0.6 1.5 0.6 1.5 0.6 1.5 0.6 0.6 1.5 0.6 0.5 0.5 0.5 0.5 0.5 0.5 0.5	95 - 99	0.5	1.0	-0.06	0.6	0.6	0.00
cale 64.2 67.8 -0.08 66.2 66.1 0.00 Acter respiratory issues 22.7 25.6 -0.07 23.9 23.8 0.00 Acter respiratory issues 1.1 1.4 -0.06 6.2 0.2 0.01 Chronic Obstructive ing disease 0.0 1.6 -0.06 2.9 2.9 0.00 Depressive disorder 2.7 4.3 -0.06 2.9 2.9 0.00 Description in the instance 2.7 4.3 -0.06 2.9 2.9 0.00 Description instance 2.7 4.3 -0.06 2.3 2.3 0.01 Gastrointestinal hesorback 2.5 2.1 -0.01 0.1 0.1 0.00 Gastrointestinal hesorback 6.2 6.3 -0.01 0.1 0.1 0.00 Main immudeficiency virus infection 0.0 0.1 0.1 0.1 0.00 0.23 0.23 0.23 0.23 0.24 0.5 0.6 0.6	100 - 104	0.1	0.1	-0.02	0.0	0.1	-0.01
Matter reprisentatory General Construction Construction	Gender: female	64.2	67.8	-0.08	66.2	66.1	0.00
Acture respiratory disease 22.7 25.6 -0.07 23.8 0.00 Attention deficit hyperstrivity disease 1.0 1.4 -0.00 0.2 0.2 0.00 Coron's disease 1.0 1.4 -0.00 0.4 0.4 0.4 0.4 0.4 0.4 0.01 0.4 0.4 0.01 0.4 0.4 0.01 0.4 0.4 0.01 0.4 0.4 0.01 0.4 0.4 0.01 0.0 0.00 0.01 0.00 0.01 0.00 0.01 0.00 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.00 0.01 0.01 0.00 0.01 0.01 0.01 0.00 0.01 0.01 0.00 0.01 0.00 0.01 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0	Medical history: General	0112	0110	0.00	0012	0011	0.00
Attention deficit hyperativity disorder 0.2 0.2 0.01 Chronic liver disease 10 1.4 -0.03 1.4 0.00 Chronic liver disease 10 1.4 -0.03 1.4 0.00 Demontis disease 2.7 4.3 -0.08 2.9 0.00 Depressive disorder 9.6 11.9 -0.07 10.1 10.0 0.00 Diabetse multitus disease 12.1 12.2 -0.01 2.3 2.3 0.00 Human immundeficiency virus infection 0.0 0.1 -0.01 7.3 2.3 0.00 Hypertensive disorder 64.6 65.2 -0.16 64.2 47.3 0.62 Distostriviter 64.6 65.8 -0.01 65.8 6.011 0.00 Destry iver 64.6 65.8 -0.02 45.2 47.3 0.62 Distostrivitis 87.9 85.2 -0.13 85.5 66.3 0.01 Destry 1.5 0.2 0.5 0.60 0.01 0.1 0.00 Restrivitis <	Acute respiratory disease	22.7	25.6	-0.07	23.9	23.8	0.00
Outronic Liver disease 1.1 1.4 -0.08 1.5 1.6 0.00 Crohis disease 0.4 0.0 0.1 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 <	Attention deficit hyperactivity disorder	0.2	0.2	0.00	0.2	0.2	0.01
Crobit's disease 1.0.4 -0.01 0.4 0.01 Demontia 2.7 4.3 -0.08 2.9 0.00 Depressive disorder 3.6 11.9 -0.07 10.1 10.0 0.00 Diabetes and tissee 12.7 2.3 0.01 0.0 0.00 Musan immundeficiency virus infection 0.0 0.1 -0.01 0.1 0.1 0.00 Hypertensive disorder 64.2 68.3 -0.01 8.5 0.01 0.00 Distorts disorder 64.4 68.3 -0.01 8.5 0.01 0.01 0.00 0.01 0.00 0.01 0.00 0.	Chronic liver disease Chronic chatructive lung disease	1.1	1.4	-0.03	1.1	1.1	0.00
Demonsion 2.7 4.3 -0.08 2.9 2.9 0.00 Depressive disorder 9.6 11.9 -0.07 10.1 10.0 0.00 Dabetes molitum 22.5 27.2 -0.11 23.7 23.5 0.00 Demonstration 0.0 0.0 -0.01 0.1 0.1 0.0 Human immundeficiency virus infection 0.0 0.0 -0.03 0.5 0.5 0.5 0.5 0.0 0.0 0.00 Hyperliphidemia 45.6 53.2 -0.13 65.5 0.5 0.00	Crohn's disease	0.4	0.4	-0.03	0.4	0.4	0.00
Depressive disorder 9.6 11.9 -0.07 10.1 10.0 0.00 Gastroesophageal reflux disease 17.1 18.5 -0.61 27.3 28.5 0.01 Buston immodef lineacy virus infection 45.6 53.2 -0.11 64.2 47.3 0.02 Hypertensive disorder 66.2 69.3 -0.11 65.0 64.5 0.01 Desity 9.4 10.6 -0.03 0.5 0.6 60.5 0.6 0.01 Desity 9.4 10.6 -0.04 8.9 8.7 0.01 Desting 9.4 10.6 -0.04 8.9 8.7 0.01 Desting 9.4 10.6 -0.04 8.7 0.7 0.00 Remaining distruction 7.0 12.9 -0.19 7.6 7.5 0.00 Rehematoid construction distase 10.7 0.2 0.2 0.2 0.2 0.00 Uicarstive colinis 0.15 0.12 0.11 0.1	Dementia	2.7	4.3	-0.08	2.9	2.9	0.00
Diabetes mollitus 22.5 27.2 -0.11 23.7 23.5 0.01 Gastrocophageal relix disease 17.1 18.6 -0.66 17.3 17.3 0.00 Gastrocophic disease 0.0 -0.15 48.2 47.3 0.00 Hyperlindenia 46.6 53.2 -0.15 48.2 47.3 0.02 Hyperlindenia 64.2 69.3 -0.11 65.0 64.5 0.01 Dateoatthritis 87.6 88.2 0.13 85.5 88.5 0.00 Descip 9.4 10.6 -0.04 8.9 8.7 0.01 Dateoatthritis 87.5 8.3.7 0.13 85.5 88.6 0.00 Shizophrenia 0.1 0.1 0.0 0.1 0.1 0.10 0.0 10.1 0.00 Uriary trainsclerous disease 11.7 13.9 -0.07 12.0 12.0 0.00 Uriary trainsclerous disease 11.7 13.9 -0.6 16.1	Depressive disorder	9.6	11.9	-0.07	10.1	10.0	0.00
unstributing if.if. 11.5 -0.00 11.3 17.0 0.01 Gasta Gasta 1.1 12.0 -0.01 1.1 0.1 0.00 Hypertensite Gasta 0.0 0.01 0.01 0.1 0.00 Hypertensite Gasta 0.01 0.01 0.1 0.00 0.00 Hypertensite Gasta 0.02 0.01 0.01 0.00 0.00 Desity 9.4 10.6 -0.04 8.9 8.7 0.01 Destarthritis 67.9 8.2 0.01 85.5 85.3 0.00 Renal inpairment 7.0 12.9 -0.02 1.5 1.5 0.00 Reconstriat 0.5 0.5 0.00 0.1 0.1 0.00 0.1 0.00 Ulcarative colitis 0.5 0.5 0.5 0.5 0.00 0.1 0.00 0.1 0.00 0.1 0.00 0.1 0.00 0.1 0.00	Diabetes mellitus	22.5	27.2	-0.11	23.7	23.5	0.01
Human immondeficiency virus infection 0.0 0.1 -0.10 0.1 0.1 0.00 Hyperingionia 45.6 53.2 -0.15 46.2 64.5 0.01 Lesion of liver 0.5 0.8 -0.03 0.5 0.5 0.00 Desity 9.4 10.6 -0.04 8.9 8.7 0.01 Detearthritis 7.9 83.2 0.13 85.5 85.3 0.00 Renal impairment 7.0 12.9 -0.19 7.6 7.5 0.00 Renal impairment 7.0 12.9 -0.19 7.6 7.5 0.00 Schizophrenia 0.1 0.1 0.00 1.0 1.0 0.00 1.0 0.00 Uriary tract infectious disease 11.7 13.9 -0.07 7.9 8 0.00 Corrobrowscular disease 11.2 12.6 -0.04 11.2 1.1 0.00 Corrobrowscular disease 13.2 2.6 -0.01 1.1 <t< td=""><td>Gastroesophageal reflux disease</td><td>17.1</td><td>19.5</td><td>-0.06</td><td>17.3</td><td>17.0</td><td>0.01</td></t<>	Gastroesophageal reflux disease	17.1	19.5	-0.06	17.3	17.0	0.01
Hyperinjedenia 45.6 53.2 -0.15 46.2 47.3 0.02 Hyperinsive disorder 64.2 69.3 -0.11 65.0 64.5 0.01 Desity 9.4 10.6 -0.04 8.9 8.7 0.01 Detexty 9.4 10.6 -0.04 8.9 8.7 0.01 Detexty 9.4 10.6 -0.04 8.9 8.7 0.01 Detexty 9.4 10.6 -0.04 8.9 8.7 0.00 Remail impairment 7.0 12.9 -0.04 3.7 3.7 0.00 Shizophrenia 0.1 0.1 0.1 0.1 0.1 0.00 0.1 0.1 0.00 Uiccrative colitis 0.5 0.6 -0.01 0.5 0.5 0.00 Uirari hepatitis C 0.2 0.3 -0.02 0.2 0.00 0 Weitagene 12.1 12.6 -0.04 11.2 11.1 0.00 Coronary arterioclerosis 19.4 21.5 -0.03 41.7 4.8 <	Human immunodeficiency virus infection	0.0	0.1	-0.01	0.1	0.1	0.00
Hypertensive disorder 64.2 69.3 -0.11 65.0 64.5 0.01 Lesion of liver 9.4 10.6 -0.03 0.5 0.5 0.00 Desity 9.4 10.6 -0.04 8.9 8.7 0.01 Deteorrhitis 87.9 83.2 0.13 85.5 85.3 0.00 Presumenta 4.6 5.6 -0.04 4.4 4.5 -0.01 Presumenta 1.5 1.7 -0.04 7.6 7.5 0.00 Schizophrenia 0.1 0.10 0.00 0.1 0.1 0.00 Urinary tract infectious disease 1.7 1.3 0.00 12.0 12.0 0.00 Verial hiprilition 9.9 12.1 -0.07 9.7 9.8 0.00 Corborary actionacular disease 11.2 12.6 -0.03 11.7 11.9 0.00 Gerebrovacular disease 9.1 9.8 -0.01 12.1 1.1 0.00	Hyperlipidemia	45.6	53.2	-0.15	48.2	47.3	0.02
Lesion of liver 0.5 0.6 -0.03 0.5 0.5 0.00 Desity 9.4 10.6 -0.04 8.9 8.7 0.01 Detoarthritis 67.9 85.2 0.14 85.5 85.3 0.01 Psoriasis 1.6 5.7 -0.04 85.5 85.3 0.01 Remal impairment 7.0 12.9 -0.19 7.6 7.5 0.00 Rheumatoid arthritis 3.5 4.3 -0.04 3.7 3.7 0.00 Uicerative colitis 0.1 0.1 0.0 0.1 0.1 0.00 Uicarthrecolitis 0.5 0.6 -0.01 1.5 0.5 0.0 Uicarthrecolitis 0.1 1.1 0.0 1.2.0 0.00 0.00 Viral hepatitis C 0.0 0.7 0.7 0.7 0.00 0.6 0.00 1.2.0 0.00 Medical history Netory 0.4 0.6 0.00 1.2.1 1.0	Hypertensive disorder	64.2	69.3	-0.11	65.0	64.5	0.01
Destary 9.4 10.6 -0.04 8.9 8.7 0.01 Deteorthritis 67.9 85.2 0.16 85.5 85.3 0.00 Persinat 4.6 5.7 -0.02 4.5 5.5 0.00 Renal impairment 7.0 12.9 -0.19 7.6 7.5 0.00 Schizophrenia 0.1 0.10 0.00 0.1 0.1 0.10 0.00 Urinary tract infectious disease 11.7 13.9 -0.07 12.0 12.0 0.00 Weical history: Cardiovascular disease 11.7 13.9 -0.07 9.7 9.8 0.00 Corobrovacular disease 11.2 11.6 -0.04 11.7 41.8 0.00 Corobrovacular disease 9.1 9.8 -0.01 1.7 41.8 0.00 Heart disease 9.1 9.8 -0.02 8.3 8.4 -0.01 Heart disease 9.1 9.8 -0.02 8.3 8.4 -0	Lesion of liver	0.5	0.8	-0.03	0.5	0.5	0.00
Drammatia its 0.4.5 0.5.6 0.4.4 0.4.5 0.0.4 Paoriasis 1.5 1.7 -0.02 1.5 1.7 -0.00 Renal impairment 7.0 12.9 -0.19 7.6 7.5 0.00 Rehumatoid arthritis 3.5 4.3 -0.04 3.7 3.7 0.00 Ulcerative colitis 0.1 0.1 0.0 0.1 0.1 0.00 Ulcarative colitis 0.5 0.6 -0.01 0.5 0.5 0.00 Viral hepatitis C 0.2 0.2 0.2 0.00 0.00 Medical history: Cardiovacular disease 1.1.2 12.6 -0.07 9.7 9.8 0.00 Coronary arteriosclerosis 19.4 21.5 -0.05 18.6 18.7 1.0 0.00 Heart disease 3.0 1.6 -0.02 8.3 8.4 -0.01 1.1 1.1 0.1 1.1 1.0 0.0 Peripheral vascular disease 9.2	Ubesity	9.4	10.6	-0.04	8.9	8.7	0.01
Provinsis 1.5 1.7 -0.02 1.5 1.5 0.00 Renal impairment 7.0 12.9 -0.19 7.6 7.5 0.00 Renumatoid arthritis 3.5 4.3 -0.44 3.7 3.7 0.00 Schizophrenia 0.5 0.6 -0.01 0.5 0.5 0.00 Urinary tract infectious disease 11.7 13.9 -0.07 12.0 0.00 Weidal history: Cardiovascular disease 11.2 12.6 -0.04 11.2 11.1 0.00 Coronary arteriosclerosis 19.4 21.5 -0.05 18.6 18.7 0.00 Heart disease 1.9 21.6 -0.06 14.7 41.8 0.00 Coronary arteriosclerosis 19.4 21.5 -0.05 18.6 18.7 0.00 Heart disease 9.1 9.8 -0.02 8.3 8.4 -0.01 Heart disease 9.2 13.6 -0.03 1.1 1.0 0.00	Pneumonia	4.6	5.6	-0.04	4.4	4.5	-0.01
Renal impairment 7.0 12.9 -0.19 7.6 7.5 0.00 Rheumatoid arthritis 3.5 4.3 -0.44 3.7 3.7 0.00 Schizophrenia 0.1 0.1 0.00 0.1 0.1 0.1 0.00 Ulcerative colltis 0.5 0.6 -0.07 12.0 12.0 0.00 Wiral hepatrits Cardiovascular disease 11.7 13.9 -0.07 12.0 12.0 0.00 Medical history: Cardiovascular disease 12.1 -0.07 9.7 9.8 0.00 Corbinovascular disease 19.4 12.6 -0.06 11.6 11.7 1.8 0.00 Heart disease 19.4 12.6 -0.06 18.6 0.00 14.8 0.00 Heart disease 9.1 9.8 -0.02 8.3 8.4 -0.01 Pulmoary emboliam 1.1 1.3 -0.01 1.1 1.0 0.00 Venous thrombosis 2.9 3.5	Psoriasis	1.5	1.7	-0.02	1.5	1.5	0.00
Rheumatoid arthritis 3.5 4.3 -0.04 3.7 3.7 0.00 Schizophrenia 0.1 0.1 0.0 0.1 0.1 0.00 Ulcerative colitis 0.5 0.6 -0.01 0.5 0.6 0.00 Viral hepatitis C 0.2 0.3 -0.02 0.2 0.2 0.00 Weideal history: Cardiovascular disease 11.2 12.6 -0.04 11.2 11.7 0.00 Gerebrovascular disease 11.2 12.6 -0.04 11.2 11.7 0.00 Gerebrovascular disease 11.2 12.6 -0.04 11.2 11.7 0.00 Heart failure 4.0 1.6 -0.03 41.7 4.6 0.00 Isochemic heart disease 9.1 9.8 -0.02 8.3 8.4 -0.01 Paripheral vascular disease 9.2 12.6 -0.01 9.7 8.7 0.00 Venous thrombosis 2.9 3.5 -0.03 3.0 3.0 0.00 Weidical history: Neoplasms 0.8 -0.02 18.8	Renal impairment	7.0	12.9	-0.19	7.6	7.5	0.00
Schlzophrenia 0.1 0.00 Viral hepatrics Cardiar disease 11.2 12.6 -0.66 16.6 16.6 16.6 16.6 6.00 16.6 -0.03 7.6 7.6 0.00 16.6 16.6 16.6 16.6 16.6 16.6 16.6 16.7 18.7 10.00 16.7 16.7 16.7 16.7 16.7 16.7 16.7 16.7 16.7 16.7 16.7 16.7 16.7 16.7 16.7	Rheumatoid arthritis	3.5	4.3	-0.04	3.7	3.7	0.00
Urinary tract infectious disease 1.7 13.9 -0.07 12.0 12.0 0.00 Wiral hepatitis 0.2 0.3 -0.02 0.2 0.00 Medical history: Cardiovascular disease 0.2 0.3 -0.07 9.7 9.8 0.00 Corenary arteriosclosis 19.4 21.5 -0.04 11.2 11.4 0.00 Heart disease 9.1 9.8 -0.02 8.8 1.0 0.00 Heart disease 9.1 9.8 -0.02 8.3 8.4 -0.01 Heart failure 8.0 1.6 -0.09 7.6 7.6 0.00 Venous thrombosis 2.9 3.5 -0.01 1.1 1.1 0.00 Venous thrombosis 2.9 3.5 -0.02 18.8 18.7 0.00 Malignant neoplasm of anoretum 0.3 0.3 0.01 0.3 0.3 0.01 Malignant tumor of breast 3.8 4.2 -0.02 4.0 4.0 0.00	Schizophrenia Ulcerative colitis	0.1	0.1	-0.01	0.1	0.1	0.00
Viral hepatitis C 0.2 0.3 -0.02 0.2 0.2 0.00 Medical history: Gardiovascular disease 11.2 12.6 -0.04 11.2 11.1 0.00 Cerebrovascular disease 19.4 21.5 -0.05 18.6 18.7 0.00 Heart disease 43.9 45.6 -0.03 41.7 41.8 0.00 Heart disease 9.1 9.8 -0.02 8.3 8.4 -0.01 Peripheral vascular disease 9.2 12.6 -0.11 9.7 0.00 Pulmoary embolism 1.1 1.3 -0.01 1.1 1.1 0.00 Wedical history: Neoplasms 2.9 3.5 -0.03 3.0 3.0 0.00 Malignant leoplasm of ancretum 0.3 0.3 0.01 0.8 0.8 0.00 Malignant tumor of foreat 3.8 4.2 -0.02 18.8 18.7 0.00 Malignant tumor of uninary bladder 0.8 0.9 0.00 0.8 0.	Urinary tract infectious disease	11.7	13.9	-0.07	12.0	12.0	0.00
Medical history: Cardiovascular disease 12.1 -0.07 9.7 9.8 0.00 Cerebrovascular disease 11.2 12.6 -0.04 11.2 11.1 0.00 Coronary arteriosclorosis 19.4 21.5 -0.05 18.6 18.7 41.8 0.00 Heart disease 43.9 45.6 -0.03 41.7 41.8 0.00 Heart disease 9.1 9.8 -0.02 8.3 8.4 -0.01 Peripheral vascular disease 9.2 12.6 -0.11 9.7 9.7 0.00 Venous thrombosis 2.9 3.5 -0.03 3.0 0.00 3.0 0.00 Malignant neoplastic disease 18.7 19.5 -0.02 8.8 0.00 3.0 0.00 Malignant neoplastic disease 18.7 19.5 -0.02 18.8 18.7 0.00 Malignant tumor of breast 3.8 4.2 -0.02 4.0 4.0 0.00 Malignant tumor of lung 0.4 <	Viral hepatitis C	0.2	0.3	-0.02	0.2	0.2	0.00
Atrial fibrillation 9.9 12. -0.07 9.7 9.8 0.00 Cerebrowscular disease 11.2 12.6 -0.04 11.2 11.1 0.00 Heart failure 8.0 10.6 -0.09 7.6 7.6 0.00 Heart failure 8.0 10.6 -0.09 7.6 7.6 0.00 Peripheral vascular disease 9.2 12.6 -0.11 9.7 9.7 0.00 Pulmonary embolism 1.1 1.3 -0.01 1.1 1.1 0.00 0.00 Malignant heoplasms 2.9 3.5 -0.03 3.0 0.00 Malignant neoplasm of anorecum 0.3 0.3 0.01 0.8 0.8 0.00 Malignant tumor of breast 3.8 4.2 -0.02 4.8 18.7 0.00 Malignant tumor of colon 0.7 0.7 0.6 0.6 0.6 0.00 Malignant tumor of lung 0.4 0.6 -0.2 0.4 0.5 -0.01 Malignant tumor of colon 0.7 7.7 0.7 0.00 <td>Medical history: Cardiovascular disease</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>	Medical history: Cardiovascular disease						
Leteblovasclini disease 11.2 12.5 -0.04 11.2 11.1 0.00 Goronary arteriosclerosis 19.4 21.5 -0.05 18.6 18.7 0.00 Heart disease 43.9 45.6 -0.03 41.7 41.8 0.00 Heart disease 9.1 9.8 -0.02 8.3 8.4 -0.01 Peripheral vascular disease 9.2 12.6 -0.11 9.7 9.7 0.00 Venous thrombosis 2.9 3.5 -0.03 3.0 3.0 0.00 Wedical history: Neoplasms -0.01 0.8 0.8 0.01 0.3 0.3 0.31 0.01 Malignant neoplasm of anorectum 0.3 0.3 0.01 0.3 0.3 0.01 Malignant tumor of breast 3.8 4.2 -0.02 4.0 4.0 0.00 Malignant tumor of lung 0.4 0.6 -0.22 0.4 0.5 -0.01 Malignant tumor of lung 0.4 0.6 -0.02 0.4 0.5 -0.01 Malignant tumor of lung 0.4	Atrial fibrillation	9.9	12.1	-0.07	9.7	9.8	0.00
Heart disease 43.9 45.6 -0.03 41.7 41.8 0.00 Heart failure 8.0 10.6 -0.09 7.6 7.6 0.00 Ischemic heart disease 9.2 12.6 -0.11 9.7 9.7 0.00 Pulmoary embolism 1.1 1.3 -0.01 1.1 1.1 0.00 0.00 Weical history: Neoplasms 2.9 3.5 -0.03 3.0 3.0 0.00 Malignant neoplasm of anorectum 0.3 0.3 0.01 0.3 0.3 0.01 Malignant tumor of breast 3.8 4.2 -0.02 4.0 4.0 0.00 Malignant tumor of colon 0.7 0.7 0.00 0.6 0.6 0.00 Malignant tumor of lung 0.4 0.6 -0.02 3.4 3.5 0.00 Malignant tumor of lung 0.4 0.6 -0.02 0.4 0.5 -0.01 Malignant tumor of lung 0.4 0.6 -0.02 0.4 0.5 -0.01 Malignant tumor of lung phosa 3.7 3.4	Coronary arteriosclerosis	19.4	21.5	-0.04	18.6	18.7	0.00
Heart failure 8.0 10.6 -0.09 7.6 7.6 0.0 Ischenic heart disease 9.1 9.8 -0.02 8.3 8.4 -0.01 Peripheral vascular disease 9.2 12.6 -0.11 9.7 9.7 0.00 Pulmonary embolism 1.1 1.3 -0.01 1.1 1.1 0.00 Wenous thrombosis 2.9 3.5 -0.03 3.0 3.0 0.00 Malignant neoplasm of anorectum 0.3 0.3 0.01 0.8 0.8 0.00 Malignant neoplasm of colon 0.7 19.5 -0.02 18.8 18.7 0.00 Malignant neoplasm of colon 0.7 0.7 0.00 0.6 0.6 0.00 Malignant tumor of lung 0.4 0.6 -0.02 0.4 0.5 -0.01 Malignant tumor of lung not prostate 3.7 3.4 0.02 3.4 3.5 0.00 Malignant tumor of systemic use 69.7 71.8 -0.05 69.	Heart disease	43.9	45.6	-0.03	41.7	41.8	0.00
Ischemic heart disease9.19.8 -0.02 8.38.4 -0.01 Peripheral vascular disease9.212.6 -0.11 9.79.70.00Pulmonary embolism1.11.3 -0.01 1.11.10.00Wenous thrombosis2.93.5 -0.03 3.03.00.00Malignant lymphoma0.81.0 -0.01 0.80.80.00Malignant neoplasm of anorectum0.30.30.010.30.30.01Malignant tumor of breast3.84.2 -0.02 18.818.70.00Malignant tumor of clon0.70.70.000.60.60.00Malignant tumor of urinary bladder0.80.90.000.80.80.00Malignant tumor of urinary bladder0.80.90.000.80.80.00Matidation use3.73.40.023.43.50.00Antidepressants29.431.7 -0.05 69.469.30.00Antidepressants29.431.7 -0.04 21.321.40.00Antiinflammatory and antirheumatic products36.935.60.0337.237.10.00Antiingragents3.025.60.1024.324.50.00Antiingrastics0.91.3 -0.07 29.329.20.00Antidepressants30.025.60.1024.324.50.00Antidepressants30.0 <td>Heart failure</td> <td>8.0</td> <td>10.6</td> <td>-0.09</td> <td>7.6</td> <td>7.6</td> <td>0.00</td>	Heart failure	8.0	10.6	-0.09	7.6	7.6	0.00
Peripheral vascular disease 9.2 12.6 -0.11 9.7 9.7 0.00 Pulmonsy embolism 1.1 1.3 -0.01 1.1 1.1 0.00 Wenous thrombosis 2.9 3.5 -0.03 3.0 0.00 Malignant neoplasm of anorectum 0.3 0.3 0.01 0.3 0.3 0.01 Malignant neoplastic disease 18.7 19.5 -0.02 18.8 18.7 0.00 Malignant tumor of colon 0.7 0.7 0.00 0.6 0.6 0.00 Malignant tumor of lung 0.4 0.6 -0.02 0.4 0.5 -0.01 Malignant tumor of urinary bladder 0.8 0.9 0.00 0.8 0.80 0.00 Medicatin use 68.7 71.8 -0.07 52.2 51.7 0.01 Antidepressants 29.4 31.7 -0.06 69.4 69.3 0.00 Antidepressants 29.4 31.7 -0.05 69.4 69.3 0.00 Antidepressants 0.9 1.3 -0.03 5.3	Ischemic heart disease	9.1	9.8	-0.02	8.3	8.4	-0.01
Yenous thrombosis 1.1 1.3 0.01 1.1 1.1 0.00 Medical history: Neoplasms Malignant lymphoma 0.8 1.0 -0.01 0.8 0.8 0.00 Malignant neoplasm of anorectum 0.3 0.3 0.01 0.3 0.3 0.01 Malignant neoplasm of anorectum 0.3 0.3 0.01 0.3 0.3 0.01 Malignant neoplasm of colon 0.7 19.5 -0.02 18.8 18.7 0.00 Malignant tumor of breast 3.8 4.2 -0.02 4.0 0.0 0.0 Malignant tumor of ol uninary bladder 0.8 0.9 0.00 0.8 0.8 0.00 Malignant tumor of urinary bladder 0.8 0.9 0.00 0.8 0.8 0.00 Medication use 65.7 71.8 -0.05 69.4 69.3 0.00 Antidepressants 29.4 31.7 -0.05 30.7 30.6 0.00 Antiinflammatory and antirheumatic products 36.9 35.6 0.03 37.2 37.1 0.00 A	Peripheral vascular disease	9.2	12.6	-0.11	9.7	9.7	0.00
Medical history: Neoplasms 0.8 0.0 0.1 <	Venous thrombosis	2.9	3.5	-0.03	3.0	3.0	0.00
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	Medical history: Neoplasms						
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	Malignant lymphoma	0.8	1.0	-0.01	0.8	0.8	0.00
Malignant tumor of breast 3.8 4.2 -0.02 4.0 4.0 0.00 Malignant tumor of colon 0.7 0.7 0.00 0.6 0.6 0.00 Malignant tumor of urinary bladder 0.8 0.9 0.00 0.8 0.00 Malignant tumor of urinary bladder 0.8 0.9 0.00 0.8 0.00 Primary malignant neoplasm of prostate 3.7 3.4 0.02 3.4 3.5 0.00 Medication use Agents acting on the renin-angiotensin system 50.6 53.9 -0.07 52.2 51.7 0.01 Antidepressants 29.4 31.7 -0.05 69.4 69.3 0.00 Antiepileptics 21.4 23.1 -0.04 21.3 21.4 0.00 Antimeoplastic agents 5.0 5.8 -0.03 5.3 5.2 0.00 Antipsoriatics 0.9 1.3 -0.03 0.9 0.9 -0.01 Antipsoriatic agents 37.4 41.8 -0.09 38.2 38.0 0.00 Calcium channel blockers 28.7	Malignant neoplasm of anorectum	0.3	0.3	0.01	0.3	0.3	0.01
Malignant tumor of colon 0.7 0.7 0.00 0.6 0.00 Malignant tumor of lung 0.4 0.6 -0.02 0.4 0.5 -0.01 Malignant tumor of urinary bladder 0.8 0.9 0.00 0.8 0.8 0.00 Primary malignant neoplasm of prostate 3.7 3.4 0.02 3.4 3.5 0.00 Matignant tumor of the renin-angiotensin system 50.6 53.9 -0.07 52.2 51.7 0.01 Antibacterials for systemic use 69.7 71.8 -0.05 69.4 69.3 0.00 Antidepressants 29.4 31.7 -0.05 30.7 30.6 0.00 Antineoplastic agents 50.0 5.8 -0.03 5.3 5.2 0.00 Antinportatics 0.9 1.3 -0.03 5.3 5.2 0.00 Antithrombotic agents 37.4 41.8 -0.09 38.2 38.0 0.00 Calcium channel blockers 28.7 31.7 -0.04 46.1 46.0 0.00 Drugs for acid related disorders 43.	Malignant neoplastic disease	18.7	19.5	-0.02	18.8	18.7	0.00
Malignant tumor of lung 0.4 0.6 -0.02 0.4 0.5 -0.01 Malignant tumor of urinary bladder 0.8 0.9 0.00 0.8 0.8 0.00 Primary malignant neoplasm of prostate 3.7 3.4 0.02 3.4 3.5 0.00 Medication use	Malignant tumor of colon	0.7	0.7	0.00	0.6	0.6	0.00
Malignant tumor of urinary bladder 0.8 0.9 0.00 0.8 0.8 0.00 Primary malignant neoplasm of prostate 3.7 3.4 0.02 3.4 3.5 0.00 Medication use 3.4 0.02 3.4 3.5 0.00 Agents acting on the renin-angiotensin system 50.6 53.9 -0.07 52.2 51.7 0.01 Antiabcterials for systemic use 69.7 71.8 -0.05 69.4 69.3 0.00 Antiepileptics 21.4 23.1 -0.04 21.3 21.4 0.00 Antiinflammatory and antirheumatic products 36.9 35.6 0.03 57.2 37.1 0.00 Antiinflammatory and antirheumatic products 36.9 1.3 -0.03 5.3 5.2 0.00 Antineoplastic agents 5.0 5.8 -0.03 5.3 5.2 0.00 Antisthrombotic agents 30.0 25.6 0.10 24.3 24.5 0.00 Beta blocking agents 37.4 41.8 -0.09 38.2 38.0 0.00 Drugs for acid re	Malignant tumor of lung	0.4	0.6	-0.02	0.4	0.5	-0.01
Primary malignant neoplasm of prostate 3.7 3.4 0.02 3.4 3.5 0.00 Medication useAgents acting on the renin-angiotensin system 50.6 53.9 -0.07 52.2 51.7 0.01 Antibacterials for systemic use 69.7 71.8 -0.05 69.4 69.3 0.00 Antidepressants 29.4 31.7 -0.05 69.4 69.3 0.00 Antiepileptics 21.4 23.1 -0.04 21.3 21.4 0.00 Antiineoplastic agents 5.0 5.8 -0.03 5.3 5.2 0.00 Antipsoriatics 0.9 1.3 -0.03 0.9 0.9 -0.01 Antiporiatics 0.9 1.3 -0.07 29.3 29.2 0.00 Calcium channel blockers 28.7 31.7 -0.07 29.3 29.2 0.00 Diuretics 46.0 47.9 -0.4 46.1 46.0 0.00 Drugs for acid related disorders 43.1 45.6 -0.05 41.4 41.7 -0.01 Drugs used in diabetes 18.4 21.7 -0.08 19.1 19.2 0.00 Immunosuppressants 4.6 6.0 -0.06 5.5 5.0 0.00 Impunse used for adhd and nootropics 2.0 2.1 -0.01 2.1 2.1 0.00	Malignant tumor of urinary bladder	0.8	0.9	0.00	0.8	0.8	0.00
Agents acting on the renin-angiotensin system 50.6 53.9 -0.07 52.2 51.7 0.01 Antibacterials for systemic use 69.7 71.8 -0.05 69.4 69.3 0.00 Antidepressants 29.4 31.7 -0.05 30.7 30.6 0.00 Antidepressants 21.4 23.1 -0.04 21.3 21.4 0.00 Antinflammatory and antirheumatic products 36.9 5.6 0.03 37.2 37.1 0.00 Antinpoplastic agents 5.0 5.8 -0.03 5.3 5.2 0.00 Antithrombotic agents 30.0 25.6 0.10 24.3 24.5 0.00 Beta blocking agents 37.4 41.8 -0.09 38.2 38.0 0.00 Calcium channel blockers 28.7 31.7 -0.07 29.3 29.2 0.00 Drugs for acid related disorders 43.1 45.6 -0.05 41.4 41.7 -0.01 Drugs for obstructive airway diseases 52.0 55.1 -0.06 5.3 0.00 0.00 Imm	Primary malignant neoplasm of prostate	3.7	3.4	0.02	3.4	3.5	0.00
Antibacterials for systemic use69.771.8-0.0569.469.30.00Antidepressants29.431.7-0.0530.730.60.00Antinflammatory and antirheumatic products36.935.60.0337.237.10.00Antineoplastic agents5.05.8-0.035.35.20.00Antineoplastic agents0.91.3-0.030.9-0.01Antineoplastic agents30.025.60.1024.324.50.00Antihrombotic agents37.441.8-0.0729.329.20.00Diuretics46.047.9-0.0446.146.00.00Drugs for acid related disorders43.145.6-0.0541.441.7-0.01Drugs for acid related disorders18.421.7-0.0819.119.20.00Immunosuppressants4.66.0-0.065.05.00.00Lipid modifying agents55.659.3-0.0857.857.20.01Drugs toracid related for adhd and nootropics2.02.1-0.0135.035.00.00Paychostimulants, agents used for adhd and nootropics2.02.1-0.012.12.10.00	Agents acting on the renin-angiotensin system	50.6	53.9	-0.07	52.2	51.7	0.01
Antidepressants29.4 31.7 -0.05 30.7 30.6 0.00 Antiepileptics 21.4 23.1 -0.04 21.3 21.4 0.00 Antiinflammatory and antirheumatic products 36.9 35.6 0.03 37.2 37.1 0.00 Antineoplastic agents 5.0 5.8 -0.03 5.3 5.2 0.00 Antiporiatics 0.9 1.3 -0.03 0.9 0.9 -0.01 Antihromotic agents 30.0 25.6 0.10 24.3 24.5 0.00 Beta blocking agents 37.4 41.8 -0.09 38.2 38.0 0.00 Calcium channel blockers 28.7 31.7 -0.07 29.3 29.2 0.00 Drugs for acid related disorders 46.0 47.9 -0.04 46.1 46.0 0.00 Drugs for obstructive airway diseases 52.0 55.1 -0.06 53.5 53.1 0.01 Drugs used in diabetes 18.4 21.7 -0.08 19.1 19.2 0.00 Immunosuppressants 4.6 6.0 -0.06 5.0 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Driods 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00 Psychostimulants, agents used for adhd and nootropics 2.0 2.1 $-$	Antibacterials for systemic use	69.7	71.8	-0.05	69.4	69.3	0.00
Antiepileptics 21.4 23.1 -0.04 21.3 21.4 0.00 Antiinflammatory and antirheumatic products 36.9 35.6 0.03 37.2 37.1 0.00 Antineoplastic agents 5.0 5.8 -0.03 5.3 5.2 0.00 Antiporiatics 0.9 1.3 -0.03 0.9 0.9 -0.01 Antithrombotic agents 30.0 25.6 0.10 24.3 24.5 0.00 Beta blocking agents 37.4 41.8 -0.09 38.2 38.0 0.00 Calcium channel blockers 28.7 31.7 -0.07 29.3 29.2 0.00 Diuretics 46.0 47.9 -0.04 46.1 46.0 0.00 Drugs for acid related disorders 43.1 45.6 -0.05 41.4 41.7 -0.01 Drugs used in diabetes 18.4 21.7 -0.06 53.5 53.1 0.01 Drugs used in diabetes 46.6 6.0 -0.06 5.0 5.0 0.00 Immunosuppressants 4.6 6.0 -0.08 57.8 57.2 0.01 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Druds 51.6 55.3 34.8 0.01 35.0 35.0 0.00 Psycholeptics 35.3 34.8 0.01 35.0 0.00 Psychostimulants, agents used for adhd and nootropics 2.0 2.1 -0.01 2.1	Antidepressants	29.4	31.7	-0.05	30.7	30.6	0.00
Antiinflammatory and antirheumatic products 36.9 35.6 0.03 37.2 37.1 0.00 Antiineoplastic agents 5.0 5.8 -0.03 5.3 5.2 0.00 Antipoplastic agents 0.9 1.3 -0.03 0.9 0.9 -0.01 Antithrombotic agents 30.0 25.6 0.10 24.3 24.5 0.00 Beta blocking agents 37.4 41.8 -0.09 38.2 38.0 0.00 Calcium channel blockers 28.7 31.7 -0.07 29.3 29.2 0.00 Diuretics 46.0 47.9 -0.04 46.1 46.0 0.00 Drugs for acid related disorders 43.1 45.6 -0.05 41.4 41.7 -0.01 Drugs for obstructive airway diseases 52.0 55.1 -0.06 53.5 53.1 0.01 Drugs used in diabetes 18.4 21.7 -0.08 19.1 19.2 0.00 Immunosuppressants 4.6 6.0 -0.06 5.0 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00	Antiepileptics	21.4	23.1	-0.04	21.3	21.4	0.00
Antibropristic agents 5.0 5.8 -0.03 5.3 5.2 -0.00 Antibropristics 0.9 1.3 -0.03 0.9 -0.01 Antibrombotic agents 30.0 25.6 0.10 24.3 24.5 0.00 Beta blocking agents 37.4 41.8 -0.09 38.2 38.0 0.00 Calcium channel blockers 28.7 31.7 -0.07 29.3 29.2 0.00 Diuretics 46.0 47.9 -0.04 46.1 46.0 0.00 Drugs for obstructive airway diseases 52.0 55.1 -0.06 53.5 53.1 0.01 Drugs used in diabetes 18.4 21.7 -0.08 19.1 19.2 0.00 Immunosuppressants 4.6 6.0 -0.06 5.0 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 </td <td>Antiinflammatory and antirheumatic products</td> <td>36.9</td> <td>35.6</td> <td>0.03</td> <td>37.2</td> <td>37.1</td> <td>0.00</td>	Antiinflammatory and antirheumatic products	36.9	35.6	0.03	37.2	37.1	0.00
Antithrombotic agents 30.0 25.6 0.10 24.3 24.5 0.00 Beta blocking agents 37.4 41.8 -0.09 38.2 38.0 0.00 Calcium channel blockers 28.7 31.7 -0.07 29.3 29.2 0.00 Diuretics 46.0 47.9 -0.04 46.1 46.0 0.00 Drugs for acid related disorders 43.1 45.6 -0.05 41.4 41.7 -0.01 Drugs for obstructive airway diseases 52.0 55.1 -0.06 53.5 53.1 0.01 Drugs used in diabetes 18.4 21.7 -0.08 19.1 19.2 0.00 Immunosuppressants 4.6 6.0 -0.06 5.0 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00 Psychostimulants, agents used for adhd and nootropics 2.0	Antipsoriatics	0.9	1.3	-0.03	0.9	0.9	-0.01
Beta blocking agents 37.4 41.8 -0.09 38.2 38.0 0.00 Calcium channel blockers 28.7 31.7 -0.07 29.3 29.2 0.00 Diuretics 46.0 47.9 -0.04 46.1 46.0 0.00 Drugs for acid related disorders 43.1 45.6 -0.05 41.4 41.7 -0.01 Drugs for obstructive airway diseases 52.0 55.1 -0.06 53.5 53.1 0.01 Drugs used in diabetes 18.4 21.7 -0.08 19.1 19.2 0.00 Immunosuppressants 4.6 6.0 -0.06 5.0 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00 Psychostimulants, agents used for adhd and nootropics 2.0 2.1 -0.01 2.1 2.1 0.00	Antithrombotic agents	30.0	25.6	0.10	24.3	24.5	0.00
Calcium channel blockers 28.7 31.7 -0.07 29.3 29.2 0.00 Diuretics 46.0 47.9 -0.04 46.1 46.0 0.00 Drugs for acid related disorders 43.1 45.6 -0.05 41.4 41.7 -0.01 Drugs for obstructive airway diseases 52.0 55.1 -0.06 53.5 53.1 0.01 Drugs used in diabetes 18.4 21.7 -0.08 19.1 19.2 0.00 Limunouppressants 4.6 6.0 -0.06 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00	Beta blocking agents	37.4	41.8	-0.09	38.2	38.0	0.00
Duretics 46.0 47.9 -0.04 46.1 46.0 0.00 Drugs for acid related disorders 43.1 45.6 -0.05 41.4 41.7 -0.01 Drugs for obstructive airway diseases 52.0 55.1 -0.06 53.5 53.1 0.01 Drugs used in diabetes 18.4 21.7 -0.08 19.1 19.2 0.00 Immunosuppressants 4.6 6.0 -0.06 5.0 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00	Calcium channel blockers	28.7	31.7	-0.07	29.3	29.2	0.00
Drugs for obstructive disorders 45.1 40.6 -0.05 41.4 41.7 -0.01 Drugs for obstructive disorders 52.0 55.1 -0.06 53.5 53.1 0.01 Drugs used in diabetes 18.4 21.7 -0.06 55.5 53.0 0.00 Immunosuppressants 4.6 6.0 -0.06 5.0 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00	Diuretics	46.0 /3 1	41.9	-0.04	40.1 /1 /	40.U /1 7	-0.00
Drugs used in diabetes 18.4 21.7 -0.08 19.1 19.2 0.00 Immunosuppressants 4.6 6.0 -0.06 5.0 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00 Psychostimulants, agents used for adhd and nootropics 2.0 2.1 -0.01 2.1 2.1 0.00	Drugs for obstructive airway diseases	52.0	55.1	-0.06	53.5	53.1	0.01
Immunosuppressants 4.6 6.0 -0.06 5.0 5.0 0.00 Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 0.00 Psychostimulants, agents used for adhd and nootropics 2.0 2.1 -0.01 2.1 2.1 0.00	Drugs used in diabetes	18.4	21.7	-0.08	19.1	19.2	0.00
Lipid modifying agents 55.6 59.3 -0.08 57.8 57.2 0.01 Opioids 51.6 45.1 0.13 46.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00 Psychostimulants, agents used for adhd and nootropics 2.0 2.1 -0.01 2.1 2.1 0.00	Immunosuppressants	4.6	6.0	-0.06	5.0	5.0	0.00
upper 51.5 45.1 0.13 45.2 46.4 0.00 Psycholeptics 35.3 34.8 0.01 35.0 35.0 0.00 Psychostimulants, agents used for adhd and nootropics 2.0 2.1 -0.01 2.1 2.1 0.00	Lipid modifying agents	55.6	59.3	-0.08	57.8	57.2	0.01
Psychostimulants, agents used for adhd and nootropics 2.0 2.1 -0.01 2.1 2.1 0.00	upioias Psycholeptics	35.3	45.1 34.8	0.13	40.∠ 35.0	40.4 35.0	0.00
	Psychostimulants, agents used for adhd and nootropic	cs 2.0	2.1	-0.01	2.1	2.1	0.00

5.6 Generalizability

The goal of any propensity score adjustments is typically to make the target and comparator cohorts comparably, to allow proper causal inference. However, in doing so, we often need to modify our population, for example dropping subjects that have no counterpart in the other exposure cohort. The population we end up estimating an effect for may end up being very different from the population we started with. An important question is: how different? And it what ways? If the populations before and after adjustment are very different, our estimated effect may not generalize to the original population (if effect modification is present). The getGeneralizabilityTable() function informs on these differences:

```
getGeneralizabilityTable(balance)
```

#	A tibble: 78	,589 x 5			
	covariateId	covariateName	beforeMatchingMean	afterMatchingMean	${\tt stdDiff}$
	<dbl></dbl>	<chr></chr>	<dbl></dbl>	<dbl></dbl>	<dbl></dbl>
1	2101660504	joint; total knee arthroplasty	0.115	0.0115	0.324
2	38003162804	s and Devices - Other Implants	0.122	0.0170	0.321
3	2105103504	cing (total knee arthroplasty)	0.112	0.0113	0.318
4	38003208804	vices - General Classification	0.135	0.0290	0.310
5	38003390804	Room - General Classification	0.130	0.0276	0.305
6	38003213804	hesia - General Classification	0.118	0.0244	0.290
7	38003138804	rmacy - General Classification	0.174	0.0668	0.283
8	38003245804	Evaluation Or Re-Evaluation	0.116	0.0289	0.272
9	2101660502	joint; total knee arthroplasty	0.168	0.0705	0.261
10	2105103502	cing (total knee arthroplasty)	0.163	0.0671	0.259
#	i 78.579 more	e rows			

In this case, because we used PS matching, we are likely aiming to estimate the average treatment effect in the treated (ATT). For this reason, the getGeneralizabilityTable() function automatically selected the target cohort as the basis for evaluating generalizability: it shows, for each covariate, the mean value before and PS adjustment in the target cohort. Also shown is the standardized difference of mean, and the table is reverse sorted by the absolute standard difference of mean (ASDM).

6 Follow-up and power

Before we start fitting an outcome model, we might be interested to know whether we have sufficient power to detect a particular effect size. It makes sense to perform these power calculations once the study population has been fully defined, so taking into account loss to the various inclusion and exclusion criteria (such as no prior outcomes), and loss due to matching and/or trimming. Since the sample size is fixed in retrospective studies (the data has already been collected), and the true effect size is unknown, the CohortMethod package provides a function to compute the minimum detectable relative risk (MDRR) instead:

```
computeMdrr(
  population = studyPop,
  modelType = "cox",
  alpha = 0.05,
  power = 0.8,
  twoSided = TRUE
)
```

targetPersons comparatorPersons targetExposures comparatorExposures targetDays comparatorDays totalOutcomes mdrr se
1 96050 155990 96050 155990 13420729 12173355 1366 1.168921 0.05571178

In this example we used the studyPop object, so the population before any matching or trimming. If we want to know the MDRR after matching, we use the matchedPop object we created earlier instead:

```
computeMdrr(
   population = matchedPop,
```

```
modelType = "cox",
alpha = 0.05,
power = 0.8,
twoSided = TRUE
)
```

targetPersons comparatorPersons targetExposures comparatorExposures targetDays comparatorDays totalOutcomes mdrr se
1 57916 57916 57916 57916 57916 8315101 4824715 649 1.246007 0.07850687

Even thought the MDRR in the matched population is higher, meaning we have less power, we should of course not be fooled: matching most likely eliminates confounding, and is therefore preferred to not matching.

To gain a better understanding of the amount of follow-up available we can also inspect the distribution of follow-up time. We defined follow-up time as time at risk, so not censored by the occurrence of the outcome. The getFollowUpDistribution can provide a simple overview:

```
getFollowUpDistribution(population = matchedPop)
```

100% 75% 50% 25% 0% Treatment
1 2 60 60 134 4373 1
2 2 45 60 74 3833 0

The output is telling us number of days of follow-up each quantile of the study population has. We can also plot the distribution:

plotFollowUpDistribution(population = matchedPop)



7 Outcome models

The outcome model is a model describing which variables are associated with the outcome.

7.1 Fitting a simple outcome model

In theory we could fit an outcome model without using the propensity scores. In this example we are fitting an outcome model using a Cox regression:

```
outcomeModel <- fitOutcomeModel(population = studyPop,</pre>
                                modelType = "cox")
outcomeModel
## Model type: cox
## Stratified: FALSE
## Use covariates: FALSE
## Use inverse probability of treatment weighting: FALSE
## Target estimand: ate
## Status: OK
##
             Estimate lower .95 upper .95
##
                                              logRr seLogRr
                                  0.98880 -0.12268 0.0569
## treatment 0.88455
                        0.79111
```

But of course we want to make use of the matching done on the propensity score:

outcomeModel

```
## Model type: cox
## Stratified: TRUE
## Use covariates: FALSE
## Use inverse probability of treatment weighting: FALSE
## Target estimand: att
## Status: OK
##
## Estimate lower .95 upper .95 logRr seLogRr
## treatment 1.010431 0.851306 1.195755 0.010377 0.0867
```

Note that we define the sub-population to be only those in the matchedPop object, which we created earlier by matching on the propensity score. We also now use a stratified Cox model, conditioning on the propensity score match sets.

Instead of matching or stratifying we can also perform Inverse Probability of Treatment Weighting (IPTW):

 ${\tt outcomeModel}$

Model type: cox
Stratified: FALSE
Use covariates: FALSE
Use inverse probability of treatment weighting: TRUE
Target estimand: att
Status: OK
##
Estimate lower .95 upper .95 logRr seLogRr
treatment 0.950147 0.765720 1.182588 -0.051139 0.1109

7.2 Adding interaction terms

We may be interested whether the effect is different across different groups in the population. To explore this, we may include interaction terms in the model. In this example we include three interaction terms:

outcomeModel

## Model type: cox ## Stratified: TRUE ## Use covariates: FALSE ## Use inverse probability of treatment weighting: FALSE ## Target estimand: ate ## Status: OK ##					
##	Estimate	lower .95	upper .95	logRr	seLogRr
## treatment	0.895362	0.717886	1.116420	-0.110528	0.1126
## treatment * gender = FEMALE	1.081270	0.865675	1.350739	0.078136	0.1135
## treatment * condition_era group during day -365 through 0 days relative to index: Type 2 diabetes mellitus	0.859767	0.673606	1.094958	-0.151094	0.1239
## treatment * drug_era group during day 0 through 0 days relative to index: ANTITHROMBOTIC AGENTS	0.792793	0.626300	1.003387	-0.232193	0.1202

Note that you can use the grepCovariateNames to find covariate IDs.

It is prudent to verify that covariate balance has also been achieved in the subgroups of interest. For example, we can check the covariate balance in the subpopulation of females:



Standardized difference of mean

7.3 Adding covariates to the outcome model

One final refinement would be to use the same covariates we used to fit the propensity model to also fit the outcome model. This way we are more robust against misspecification of the model, and more likely to remove bias. For this we use the regularized Cox regression in the Cyclops package. (Note that the treatment variable is automatically excluded from regularization.)

outcomeModel

```
## Model type: cox
## Stratified: TRUE
## Use covariates: TRUE
## Use inverse probability of treatment weighting: FALSE
## Target estimand: att
## Status: OK
## Prior variance: 0.0075309854039699
##
## Estimate lower .95 upper .95 logRr seLogRr
## treatment 0.964765 0.796861 1.163935 -0.035871 0.0967
```

7.4 Inspecting the outcome model

We can inspect more details of the outcome model:

```
exp(coef(outcomeModel))
```

900000010805
0.9647651
exp(confint(outcomeModel))

```
## [1] 0.7968613 1.1639351
```

We can also see the covariates that ended up in the outcome model:

getOutcomeModel(outcomeModel, cohortMethodData)

coefficient id name
1 -0.03587062 9e+11 Treatment

7.5 Kaplan-Meier plot

We can create the Kaplan-Meier plot:

plotKaplanMeier(matchedPop, includeZero = FALSE)



Note that the Kaplan-Meier plot will automatically adjust for any stratification, matching, or trimming that may have been applied.

7.6 Time-to-event plot

We can also plot time-to-event, showing both events before and after the index date, and events during and outside the defined time-at-risk window. This plot can provide insight into the temporal pattern of the outcome relative to the exposures:



Note that this plot does not show any adjustment for the propensity score.

8 Acknowledgments

Considerable work has been dedicated to provide the CohortMethod package.

```
citation("CohortMethod")
```

```
## To cite package 'CohortMethod' in publications use:
##
     Schuemie M, Suchard M, Ryan P (2024). _CohortMethod: New-User Cohort Method with Large Scale Proper
##
     https://github.com/OHDSI/CohortMethod.
##
##
## A BibTeX entry for LaTeX users is
##
##
     @Manual{,
       title = {CohortMethod: New-User Cohort Method with Large Scale Propensity and Outcome
##
## Models},
       author = {Martijn Schuemie and Marc Suchard and Patrick Ryan},
##
##
       year = \{2024\},
##
       note = {https://ohdsi.github.io/CohortMethod,
## https://github.com/OHDSI/CohortMethod},
##
     }
```

Further, CohortMethod makes extensive use of the Cyclops package.

citation("Cyclops")

```
## To cite Cyclops in publications use:
##
     Suchard MA, Simpson SE, Zorych I, Ryan P, Madigan D (2013). "Massive parallelization of serial inf
##
     Modeling and Computer Simulation_, *23*, 10. <https://dl.acm.org/doi/10.1145/2414416.2414791>.
##
##
## A BibTeX entry for LaTeX users is
##
     @Article{,
##
##
       author = {M. A. Suchard and S. E. Simpson and I. Zorych and P. Ryan and D. Madigan},
##
       title = {Massive parallelization of serial inference algorithms for complex generalized linear m
##
       journal = {ACM Transactions on Modeling and Computer Simulation},
       volume = \{23\},
##
##
       pages = \{10\},\
##
       year = \{2013\},\
##
       url = {https://dl.acm.org/doi/10.1145/2414416.2414791},
##
     }
```

This work is supported in part through the National Science Foundation grant IIS 1251151.