R version 3.6.3 (2020-02-29) -- "Holding the Windsock"

Copyright (C) 2020 The R Foundation for Statistical Computing

Platform: x86\_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.

You are welcome to redistribute it under certain conditions.

Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.

Type 'contributors()' for more information and

'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or

'help.start()' for an HTML browser interface to help.

Type 'q()' to quit R.

[Previously saved workspace restored]

> # rm(list=ls())

> ## Please read these comments before running the code contained in this R script!

> ##

> ## In order for the code to execute, users need to set the object: dir\_supplemental (line 133)

> ## to be the parent directory of the "code" folder contained in the supplemental material.

> ## For additional details regarding the analyses performed here, please refer to the manuscript.

> ##

> ## The code below is divided into the following "chunks"

> ## - Section 0: Install and load all necessary packages including the rnhanesdata package

> ## which contains the accelerometry data used in our analysis.

> ## - Section 1:

> ## 1a. Load and merge accelerometry, demographic/comorbidity, and mortality data

> ## 1b. Create new factor variables which we will use in our analysis.

> ## \* We collapse most comorbidity data into either "Yes", "No", or msising from NHANES

> ## questionairre data which allows for "don't know" or "refused" as responses.

> ## We assume individuals who respond "don't know" or "missing" do not have that particular condition.

> ## \* We collapse adult education into 3 levels: "less than high school", "high school", and "more than highschool"

> ## \* We add a "missing" level to alcohol consumption in order to retain individuals with missing data

> ## for this item in our analysis.

> ##

> ## - Section 2: Calculate commonly used accelerometry summary measures:

> ## \* TAC: Total activity counts

> ## \* TLAC: Total log(1+activity countss)

> ## \* WT: Total wear time

> ## \* ST: Total sedentary, sleep, or non-wear time

> ## \* MVPA: Total time spent in MVPA

> ## \* SATP: Transition probability from sedentary, sleep, or non-wear to active. In the manuscript this is referred to

> ## using the subscript sl/nw

> ## \* ASTP: Transition probability from active to sedentary, sleep, or non-wear. In the manuscript this is referred to

> ## using the subscript sl/nw

> ## Note that in this section these variables are calculated at the day level. After applying exlcusion criteria

> ## in Section 3, we average across days within individuals to get one number per measure for each participant.

> ## The exclusion criteria involves excluding days which are deemed to have insufficient wear-time (<10 hours)

> ##

> ## - Section 3: Apply exclusion criteria and create a data frame with one row per subject which will be used as a basis for regression

> ## analyses. This data frame is called "data\_analysis".

> ##

> ## When estimating complex survey generalized linear models, we create a svydesign() object via the survey package which

> ## uses "data\_analysis". In order to fit models which use both the "adjusted" or "unadjusted" survey weights, we create two separate

> ## svydesign() objects. This is done in Section 4.d when we perform forward selection.

> ##

> ## Once we've subset the data to obtain "data\_analysis", we calculate both adjusted and unadjsuted normalized

> ## survey weights. These weights are calculated using the reweight\_accel() function

> ## (see ?reweight\_accel for details). The adjusted weights are calculated using age, gender,

> ## and ethnicity strata. The "adjusted" normalizaed weights we use for regression analyses are "wtmec4yr\_adj\_norm"

> ##

> ## Ecxlusion criteria

> ## \* Apply age exclusion (i.e. younger than 50, or 85 and over). Note that individuals age 85 and over

> ## at the time of accelerometer wear have NA (missing) values for the variable RIDAGEEX which records

> ## age in months at the time individuals took part in the exam portion of the study.

> ## \* Create a table of pairwise missing data on variables that we intend to include in our prediction model to

> ## see the distribution of missing data. In our analysis individuals are only excluded for:

> ## - Missing body mass index (BMI)

> ## - Missing education

> ## - "Bad" accelerometry data

> ## + fewer than 3 days of accelerometry data with at least 10 hours of estimated wear-time

> ## + device calibration flag recorded by NHANES

> ## + data reliability flag recorded by NHANES

> ## - Missing mortality data

> ## - Individuals recorded as "alive" but had fewer than 5 years of follow-up

> ##

> ## - Section 4: Data analysis

> ## 4a. Perform (unweighted) functional principal component analysis (FPCA) and survey weighted PCA

> ## 4b. Use backward selection to identify FPCA features associated with 5-year mortality. Find surrogate measures on

> ## the log transformed activity counts that correlate strongly with the features which are associated with 5-year mortality.

> ## 4c. Perform scalar on function regression (SoFR) where we include individuals' average (log-transformed) activity profiles

> ## as the functional predictor.

> ## 4d. Use forward selection to evaluate the predictive value of our set of variables identified in 4b/4c as well as the commonly used

> ## accelerometry features calculated in Section 2 of this code.

> ##

> ## \* Note that functional forms are assumed to be linear and no interactions/effect modifications are considered.

> ##

> ##

> ##

> ## By default, the assummption is that this R script will be downloaded with the supplemental material which accompanies

> ## the manuscript "Organizing and analyzing the activity in NHANES". This supplemental material is a zipped file with 3 folders:

> ## - code

> ## - figures

> ## - tables

> ## Assuming users correctly set the variable "dir\_supplemental",

> ## running this script will save all figure output to the "figures" folder and

> ## .tex files which contain the latex version of tables presented in the manuscript.

> ##

> ## Because the data are relatively large and require a non-trivial amount of working memory (RAM)

> ## we "clean up" the workspace as we go, clearing items after the relevant figures/tables/regression results are

> ## created/printed to the console. This may need to be taken into consideration if there's a particualr area you'd like to investigate in more detail.

> ##

> ## Finally, although most of the code executes fairly quickly, there are two sections that may take a few minutes to run.

> ## The first, calculating survey weighted principal components, will not execute by default. This can be changed by switching

> ## make\_plot\_fpca\_vs\_svypca to TRUE in Section 0.

> ## The other chunk that takes time to run is the forward selection procedure. There's a print statement embedded in the code that will

> ## report progress on the procedure, but this may take 15-25 minutes to completely finish.

>

>

> ########################################

> ## ##

> ## Section 0: load required packages ##

> ## ##

> ########################################

>

>

> ## Check for packages needed to run analyses/install the rnhanesdata package.

> ## Note: all these packages are available on CRAN and can therefore be downloaded using the install.packages() function,

> ## the rnhanesdata package is not on CRAN due to package size

> pckgs <- c("tableone","knitr","kableExtra", ## packages used for creating Table 1

+ "devtools", ## package used to download R packages stored on GitHub

+ "magrittr","dplyr", ## packages for merging/transforming data

+ "survey", ## package used for analyzing complex survey data in R

+ "mgcv","refund" ## packages used for smoothing/functional regression

+ )

>

> sapply(pckgs, function(x) if(!require(x,character.only=TRUE,quietly=TRUE)) {

+ install.packages(x)

+ require(x, character.only=TRUE)

+ })

Installing package into ‘C:/Users/adlik/OneDrive/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

--- Please select a CRAN mirror for use in this session ---

Warning: dependency ‘survey’ is not available

also installing the dependencies ‘bit’, ‘bit64’, ‘progress’, ‘vroom’, ‘tzdb’, ‘fansi’, ‘utf8’, ‘gtools’, ‘cli’, ‘forcats’, ‘hms’, ‘readr’, ‘tibble’, ‘glue’, ‘magrittr’, ‘pillar’, ‘stringi’, ‘purrr’, ‘proxy’, ‘gdata’, ‘haven’, ‘dplyr’, ‘lifecycle’, ‘rlang’, ‘vctrs’, ‘stringr’, ‘tidyr’, ‘tidyselect’, ‘e1071’, ‘zoo’, ‘gmodels’, ‘labelled’

Warning: unable to access index for repository https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6:

cannot open URL 'https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6/PACKAGES'

Packages which are only available in source form, and may need

compilation of C/C++/Fortran: ‘bit’ ‘bit64’ ‘vroom’ ‘tzdb’ ‘fansi’

‘utf8’ ‘gtools’ ‘cli’ ‘readr’ ‘tibble’ ‘glue’ ‘magrittr’ ‘stringi’

‘purrr’ ‘proxy’ ‘haven’ ‘dplyr’ ‘rlang’ ‘vctrs’ ‘tidyr’ ‘tidyselect’

‘e1071’ ‘zoo’

These will not be installed

installing the source packages ‘progress’, ‘forcats’, ‘hms’, ‘pillar’, ‘gdata’, ‘lifecycle’, ‘stringr’, ‘gmodels’, ‘labelled’, ‘tableone’

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/progress\_1.2.3.tar.gz'

Content type 'application/x-gzip' length 30499 bytes (29 KB)

downloaded 29 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/forcats\_1.0.0.tar.gz'

Content type 'application/x-gzip' length 294183 bytes (287 KB)

downloaded 287 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/hms\_1.1.3.tar.gz'

Content type 'application/x-gzip' length 43379 bytes (42 KB)

downloaded 42 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/pillar\_1.9.0.tar.gz'

Content type 'application/x-gzip' length 444528 bytes (434 KB)

downloaded 434 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/gdata\_3.0.0.tar.gz'

Content type 'application/x-gzip' length 326299 bytes (318 KB)

downloaded 318 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/lifecycle\_1.0.4.tar.gz'

Content type 'application/x-gzip' length 107656 bytes (105 KB)

downloaded 105 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/stringr\_1.5.1.tar.gz'

Content type 'application/x-gzip' length 176599 bytes (172 KB)

downloaded 172 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/gmodels\_2.19.1.tar.gz'

Content type 'application/x-gzip' length 42102 bytes (41 KB)

downloaded 41 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/labelled\_2.13.0.tar.gz'

Content type 'application/x-gzip' length 240347 bytes (234 KB)

downloaded 234 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/tableone\_0.13.2.tar.gz'

Content type 'application/x-gzip' length 264760 bytes (258 KB)

downloaded 258 KB

ERROR: dependency 'gtools' is not available for package 'gdata'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/gdata'

ERROR: dependencies 'cli', 'glue', 'rlang' are not available for package 'lifecycle'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/lifecycle'

ERROR: dependencies 'cli', 'glue', 'lifecycle', 'magrittr', 'rlang', 'tibble' are not available for package 'forcats'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/forcats'

ERROR: dependencies 'lifecycle', 'rlang', 'vctrs' are not available for package 'hms'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/hms'

ERROR: dependencies 'cli', 'fansi', 'glue', 'lifecycle', 'rlang', 'utf8', 'vctrs' are not available for package 'pillar'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/pillar'

ERROR: dependencies 'cli', 'glue', 'lifecycle', 'magrittr', 'rlang', 'stringi', 'vctrs' are not available for package 'stringr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/stringr'

ERROR: dependency 'gdata' is not available for package 'gmodels'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/gmodels'

ERROR: dependency 'hms' is not available for package 'progress'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/progress'

ERROR: dependencies 'haven', 'dplyr', 'lifecycle', 'rlang', 'vctrs', 'stringr', 'tidyr', 'tidyselect' are not available for package 'labelled'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/labelled'

ERROR: dependencies 'survey', 'e1071', 'zoo', 'gmodels', 'labelled' are not available for package 'tableone'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/tableone'

The downloaded source packages are in

‘C:\Users\adlik\AppData\Local\Temp\RtmpWe4OTi\downloaded\_packages’

Loading required package: tableone

Installing package into ‘C:/Users/adlik/OneDrive/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

Warning: dependency ‘evaluate’ is not available

also installing the dependencies ‘highr’, ‘xfun’, ‘yaml’

Warning: unable to access index for repository https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6:

cannot open URL 'https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6/PACKAGES'

Packages which are only available in source form, and may need

compilation of C/C++/Fortran: ‘xfun’ ‘yaml’

These will not be installed

installing the source packages ‘highr’, ‘knitr’

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/highr\_0.11.tar.gz'

Content type 'application/x-gzip' length 13851 bytes (13 KB)

downloaded 13 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/knitr\_1.47.tar.gz'

Content type 'application/x-gzip' length 587962 bytes (574 KB)

downloaded 574 KB

ERROR: dependency 'xfun' is not available for package 'highr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/highr'

ERROR: dependencies 'evaluate', 'highr', 'xfun', 'yaml' are not available for package 'knitr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/knitr'

The downloaded source packages are in

‘C:\Users\adlik\AppData\Local\Temp\RtmpWe4OTi\downloaded\_packages’

Loading required package: knitr

Installing package into ‘C:/Users/adlik/OneDrive/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

Warning: dependency ‘evaluate’ is not available

also installing the dependencies ‘fs’, ‘rappdirs’, ‘cachem’, ‘memoise’, ‘mime’, ‘sass’, ‘colorspace’, ‘highr’, ‘xfun’, ‘yaml’, ‘cli’, ‘glue’, ‘lifecycle’, ‘rlang’, ‘stringi’, ‘vctrs’, ‘bslib’, ‘fontawesome’, ‘jquerylib’, ‘jsonlite’, ‘tinytex’, ‘farver’, ‘munsell’, ‘base64enc’, ‘fastmap’, ‘systemfonts’, ‘knitr’, ‘magrittr’, ‘stringr’, ‘xml2’, ‘rmarkdown’, ‘scales’, ‘htmltools’, ‘digest’, ‘svglite’

Warning: unable to access index for repository https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6:

cannot open URL 'https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6/PACKAGES'

Packages which are only available in source form, and may need

compilation of C/C++/Fortran: ‘fs’ ‘rappdirs’ ‘cachem’ ‘mime’ ‘sass’

‘colorspace’ ‘xfun’ ‘yaml’ ‘cli’ ‘glue’ ‘rlang’ ‘stringi’ ‘vctrs’

‘jsonlite’ ‘farver’ ‘base64enc’ ‘fastmap’ ‘systemfonts’ ‘magrittr’

‘xml2’ ‘scales’ ‘htmltools’ ‘digest’ ‘svglite’

These will not be installed

installing the source packages ‘memoise’, ‘highr’, ‘lifecycle’, ‘bslib’, ‘fontawesome’, ‘jquerylib’, ‘tinytex’, ‘munsell’, ‘knitr’, ‘stringr’, ‘rmarkdown’, ‘kableExtra’

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/memoise\_2.0.1.tar.gz'

Content type 'application/x-gzip' length 17852 bytes (17 KB)

downloaded 17 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/highr\_0.11.tar.gz'

Content type 'application/x-gzip' length 13851 bytes (13 KB)

downloaded 13 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/lifecycle\_1.0.4.tar.gz'

Content type 'application/x-gzip' length 107656 bytes (105 KB)

downloaded 105 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/bslib\_0.7.0.tar.gz'

Content type 'application/x-gzip' length 5125401 bytes (4.9 MB)

downloaded 4.9 MB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/fontawesome\_0.5.2.tar.gz'

Content type 'application/x-gzip' length 1281884 bytes (1.2 MB)

downloaded 1.2 MB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/jquerylib\_0.1.4.tar.gz'

Content type 'application/x-gzip' length 520207 bytes (508 KB)

downloaded 508 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/tinytex\_0.51.tar.gz'

Content type 'application/x-gzip' length 34491 bytes (33 KB)

downloaded 33 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/munsell\_0.5.1.tar.gz'

Content type 'application/x-gzip' length 182310 bytes (178 KB)

downloaded 178 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/knitr\_1.47.tar.gz'

Content type 'application/x-gzip' length 587962 bytes (574 KB)

downloaded 574 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/stringr\_1.5.1.tar.gz'

Content type 'application/x-gzip' length 176599 bytes (172 KB)

downloaded 172 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/rmarkdown\_2.27.tar.gz'

Content type 'application/x-gzip' length 2190416 bytes (2.1 MB)

downloaded 2.1 MB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/kableExtra\_1.4.0.tar.gz'

Content type 'application/x-gzip' length 1824636 bytes (1.7 MB)

downloaded 1.7 MB

ERROR: dependencies 'rlang', 'cachem' are not available for package 'memoise'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/memoise'

ERROR: dependency 'xfun' is not available for package 'highr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/highr'

ERROR: dependencies 'cli', 'glue', 'rlang' are not available for package 'lifecycle'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/lifecycle'

ERROR: dependencies 'rlang', 'htmltools' are not available for package 'fontawesome'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/fontawesome'

ERROR: dependency 'htmltools' is not available for package 'jquerylib'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/jquerylib'

ERROR: dependency 'xfun' is not available for package 'tinytex'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/tinytex'

ERROR: dependency 'colorspace' is not available for package 'munsell'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/munsell'

ERROR: dependencies 'base64enc', 'cachem', 'fastmap', 'htmltools', 'jquerylib', 'jsonlite', 'lifecycle', 'memoise', 'mime', 'rlang', 'sass' are not available for package 'bslib'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/bslib'

ERROR: dependencies 'evaluate', 'highr', 'xfun', 'yaml' are not available for package 'knitr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/knitr'

ERROR: dependencies 'cli', 'glue', 'lifecycle', 'magrittr', 'rlang', 'stringi', 'vctrs' are not available for package 'stringr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/stringr'

ERROR: dependencies 'bslib', 'evaluate', 'fontawesome', 'htmltools', 'jquerylib', 'jsonlite', 'knitr', 'tinytex', 'xfun', 'yaml' are not available for package 'rmarkdown'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/rmarkdown'

ERROR: dependencies 'knitr', 'magrittr', 'stringr', 'xml2', 'rmarkdown', 'scales', 'htmltools', 'digest', 'svglite' are not available for package 'kableExtra'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/kableExtra'

The downloaded source packages are in

‘C:\Users\adlik\AppData\Local\Temp\RtmpWe4OTi\downloaded\_packages’

Loading required package: kableExtra

Installing package into ‘C:/Users/adlik/OneDrive/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

Warning: dependencies ‘downlit’, ‘evaluate’ are not available

also installing the dependencies ‘Rcpp’, ‘utf8’, ‘askpass’, ‘credentials’, ‘openssl’, ‘sys’, ‘zip’, ‘httr2’, ‘fastmap’, ‘httpuv’, ‘mime’, ‘fontawesome’, ‘sourcetools’, ‘later’, ‘promises’, ‘base64enc’, ‘jquerylib’, ‘sass’, ‘systemfonts’, ‘textshaping’, ‘tinytex’, ‘xfun’, ‘fansi’, ‘pillar’, ‘highr’, ‘diffobj’, ‘rematch2’, ‘curl’, ‘gert’, ‘gh’, ‘glue’, ‘jsonlite’, ‘purrr’, ‘rappdirs’, ‘yaml’, ‘cachem’, ‘shiny’, ‘htmltools’, ‘callr’, ‘processx’, ‘bslib’, ‘digest’, ‘httr’, ‘magrittr’, ‘ragg’, ‘rmarkdown’, ‘tibble’, ‘xml2’, ‘htmlwidgets’, ‘stringr’, ‘vctrs’, ‘xopen’, ‘commonmark’, ‘knitr’, ‘stringi’, ‘brio’, ‘ps’, ‘waldo’, ‘usethis’, ‘cli’, ‘desc’, ‘ellipsis’, ‘fs’, ‘lifecycle’, ‘memoise’, ‘miniUI’, ‘pkgbuild’, ‘pkgdown’, ‘pkgload’, ‘profvis’, ‘rcmdcheck’, ‘rlang’, ‘roxygen2’, ‘rversions’, ‘sessioninfo’, ‘testthat’, ‘urlchecker’

Warning: unable to access index for repository https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6:

cannot open URL 'https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6/PACKAGES'

Packages which are only available in source form, and may need

compilation of C/C++/Fortran: ‘Rcpp’ ‘utf8’ ‘askpass’ ‘openssl’ ‘sys’

‘zip’ ‘fastmap’ ‘httpuv’ ‘mime’ ‘sourcetools’ ‘later’ ‘promises’

‘base64enc’ ‘sass’ ‘systemfonts’ ‘textshaping’ ‘xfun’ ‘fansi’

‘diffobj’ ‘curl’ ‘gert’ ‘glue’ ‘jsonlite’ ‘purrr’ ‘rappdirs’ ‘yaml’

‘cachem’ ‘htmltools’ ‘processx’ ‘digest’ ‘magrittr’ ‘ragg’ ‘tibble’

‘xml2’ ‘vctrs’ ‘commonmark’ ‘stringi’ ‘brio’ ‘ps’ ‘cli’ ‘ellipsis’

‘fs’ ‘profvis’ ‘rlang’ ‘roxygen2’ ‘testthat’

These will not be installed

installing the source packages ‘credentials’, ‘httr2’, ‘fontawesome’, ‘jquerylib’, ‘tinytex’, ‘pillar’, ‘highr’, ‘rematch2’, ‘gh’, ‘shiny’, ‘callr’, ‘bslib’, ‘httr’, ‘rmarkdown’, ‘htmlwidgets’, ‘stringr’, ‘xopen’, ‘knitr’, ‘waldo’, ‘usethis’, ‘desc’, ‘lifecycle’, ‘memoise’, ‘miniUI’, ‘pkgbuild’, ‘pkgdown’, ‘pkgload’, ‘rcmdcheck’, ‘rversions’, ‘sessioninfo’, ‘urlchecker’, ‘devtools’

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/credentials\_2.0.1.tar.gz'

Content type 'application/x-gzip' length 283844 bytes (277 KB)

downloaded 277 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/httr2\_1.0.1.tar.gz'

Content type 'application/x-gzip' length 214736 bytes (209 KB)

downloaded 209 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/fontawesome\_0.5.2.tar.gz'

Content type 'application/x-gzip' length 1281884 bytes (1.2 MB)

downloaded 1.2 MB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/jquerylib\_0.1.4.tar.gz'

Content type 'application/x-gzip' length 520207 bytes (508 KB)

downloaded 508 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/tinytex\_0.51.tar.gz'

Content type 'application/x-gzip' length 34491 bytes (33 KB)

downloaded 33 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/pillar\_1.9.0.tar.gz'

Content type 'application/x-gzip' length 444528 bytes (434 KB)

downloaded 434 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/highr\_0.11.tar.gz'

Content type 'application/x-gzip' length 13851 bytes (13 KB)

downloaded 13 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/rematch2\_2.1.2.tar.gz'

Content type 'application/x-gzip' length 13366 bytes (13 KB)

downloaded 13 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/gh\_1.4.1.tar.gz'

Content type 'application/x-gzip' length 46164 bytes (45 KB)

downloaded 45 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/shiny\_1.8.1.1.tar.gz'

Content type 'application/x-gzip' length 3049516 bytes (2.9 MB)

downloaded 2.9 MB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/callr\_3.7.6.tar.gz'

Content type 'application/x-gzip' length 104358 bytes (101 KB)

downloaded 101 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/bslib\_0.7.0.tar.gz'

Content type 'application/x-gzip' length 5125401 bytes (4.9 MB)

downloaded 4.9 MB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/httr\_1.4.7.tar.gz'

Content type 'application/x-gzip' length 118500 bytes (115 KB)

downloaded 115 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/rmarkdown\_2.27.tar.gz'

Content type 'application/x-gzip' length 2190416 bytes (2.1 MB)

downloaded 2.1 MB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/htmlwidgets\_1.6.4.tar.gz'

Content type 'application/x-gzip' length 868895 bytes (848 KB)

downloaded 848 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/stringr\_1.5.1.tar.gz'

Content type 'application/x-gzip' length 176599 bytes (172 KB)

downloaded 172 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/xopen\_1.0.1.tar.gz'

Content type 'application/x-gzip' length 11855 bytes (11 KB)

downloaded 11 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/knitr\_1.47.tar.gz'

Content type 'application/x-gzip' length 587962 bytes (574 KB)

downloaded 574 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/waldo\_0.5.2.tar.gz'

Content type 'application/x-gzip' length 39711 bytes (38 KB)

downloaded 38 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/usethis\_2.2.3.tar.gz'

Content type 'application/x-gzip' length 371088 bytes (362 KB)

downloaded 362 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/desc\_1.4.3.tar.gz'

Content type 'application/x-gzip' length 80066 bytes (78 KB)

downloaded 78 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/lifecycle\_1.0.4.tar.gz'

Content type 'application/x-gzip' length 107656 bytes (105 KB)

downloaded 105 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/memoise\_2.0.1.tar.gz'

Content type 'application/x-gzip' length 17852 bytes (17 KB)

downloaded 17 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/miniUI\_0.1.1.1.tar.gz'

Content type 'application/x-gzip' length 97958 bytes (95 KB)

downloaded 95 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/pkgbuild\_1.4.4.tar.gz'

Content type 'application/x-gzip' length 48794 bytes (47 KB)

downloaded 47 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/pkgdown\_2.0.9.tar.gz'

Content type 'application/x-gzip' length 855137 bytes (835 KB)

downloaded 835 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/pkgload\_1.3.4.tar.gz'

Content type 'application/x-gzip' length 78189 bytes (76 KB)

downloaded 76 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/rcmdcheck\_1.4.0.tar.gz'

Content type 'application/x-gzip' length 73522 bytes (71 KB)

downloaded 71 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/rversions\_2.1.2.tar.gz'

Content type 'application/x-gzip' length 42135 bytes (41 KB)

downloaded 41 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/sessioninfo\_1.2.2.tar.gz'

Content type 'application/x-gzip' length 170924 bytes (166 KB)

downloaded 166 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/urlchecker\_1.0.1.tar.gz'

Content type 'application/x-gzip' length 13340 bytes (13 KB)

downloaded 13 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/devtools\_2.4.5.tar.gz'

Content type 'application/x-gzip' length 374718 bytes (365 KB)

downloaded 365 KB

ERROR: dependencies 'openssl', 'sys', 'curl', 'jsonlite', 'askpass' are not available for package 'credentials'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/credentials'

ERROR: dependencies 'rlang', 'htmltools' are not available for package 'fontawesome'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/fontawesome'

ERROR: dependency 'htmltools' is not available for package 'jquerylib'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/jquerylib'

ERROR: dependency 'xfun' is not available for package 'tinytex'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/tinytex'

ERROR: dependency 'xfun' is not available for package 'highr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/highr'

ERROR: dependency 'tibble' is not available for package 'rematch2'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/rematch2'

ERROR: dependency 'processx' is not available for package 'callr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/callr'

ERROR: dependencies 'curl', 'jsonlite', 'mime', 'openssl' are not available for package 'httr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/httr'

ERROR: dependency 'processx' is not available for package 'xopen'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/xopen'

ERROR: dependency 'cli' is not available for package 'desc'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/desc'

ERROR: dependencies 'cli', 'glue', 'rlang' are not available for package 'lifecycle'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/lifecycle'

ERROR: dependencies 'rlang', 'cachem' are not available for package 'memoise'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/memoise'

ERROR: dependencies 'curl', 'xml2' are not available for package 'rversions'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/rversions'

ERROR: dependency 'cli' is not available for package 'sessioninfo'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/sessioninfo'

ERROR: dependencies 'cli', 'curl', 'xml2' are not available for package 'urlchecker'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/urlchecker'

ERROR: dependencies 'cli', 'curl', 'glue', 'lifecycle', 'magrittr', 'openssl', 'rappdirs', 'rlang', 'vctrs' are not available for package 'httr2'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/httr2'

ERROR: dependencies 'cli', 'fansi', 'glue', 'lifecycle', 'rlang', 'utf8', 'vctrs' are not available for package 'pillar'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/pillar'

ERROR: dependencies 'base64enc', 'cachem', 'fastmap', 'htmltools', 'jquerylib', 'jsonlite', 'lifecycle', 'memoise', 'mime', 'rlang', 'sass' are not available for package 'bslib'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/bslib'

ERROR: dependencies 'cli', 'glue', 'lifecycle', 'magrittr', 'rlang', 'stringi', 'vctrs' are not available for package 'stringr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/stringr'

ERROR: dependencies 'evaluate', 'highr', 'xfun', 'yaml' are not available for package 'knitr'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/knitr'

ERROR: dependencies 'cli', 'diffobj', 'fansi', 'glue', 'rematch2', 'rlang', 'tibble' are not available for package 'waldo'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/waldo'

ERROR: dependencies 'callr', 'cli', 'desc', 'processx' are not available for package 'pkgbuild'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/pkgbuild'

ERROR: dependencies 'cli', 'glue', 'httr2', 'jsonlite', 'lifecycle', 'rlang' are not available for package 'gh'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/gh'

ERROR: dependencies 'httpuv', 'mime', 'jsonlite', 'fontawesome', 'htmltools', 'sourcetools', 'later', 'promises', 'rlang', 'fastmap', 'commonmark', 'glue', 'bslib', 'cachem', 'lifecycle' are not available for package 'shiny'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/shiny'

ERROR: dependencies 'bslib', 'evaluate', 'fontawesome', 'htmltools', 'jquerylib', 'jsonlite', 'knitr', 'tinytex', 'xfun', 'yaml' are not available for package 'rmarkdown'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/rmarkdown'

ERROR: dependencies 'cli', 'desc', 'fs', 'glue', 'pkgbuild', 'rlang' are not available for package 'pkgload'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/pkgload'

ERROR: dependencies 'callr', 'cli', 'curl', 'desc', 'digest', 'pkgbuild', 'sessioninfo', 'xopen' are not available for package 'rcmdcheck'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/rcmdcheck'

ERROR: dependencies 'htmltools', 'jsonlite', 'knitr', 'rmarkdown', 'yaml' are not available for package 'htmlwidgets'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/htmlwidgets'

ERROR: dependencies 'cli', 'curl', 'desc', 'fs', 'gert', 'gh', 'glue', 'jsonlite', 'lifecycle', 'purrr', 'rappdirs', 'rlang', 'yaml' are not available for package 'usethis'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/usethis'

ERROR: dependencies 'shiny', 'htmltools' are not available for package 'miniUI'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/miniUI'

ERROR: dependencies 'bslib', 'callr', 'cli', 'desc', 'digest', 'downlit', 'fs', 'httr', 'jsonlite', 'magrittr', 'memoise', 'purrr', 'ragg', 'rlang', 'rmarkdown', 'tibble', 'xml2', 'yaml' are not available for package 'pkgdown'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/pkgdown'

ERROR: dependencies 'usethis', 'cli', 'desc', 'ellipsis', 'fs', 'lifecycle', 'memoise', 'miniUI', 'pkgbuild', 'pkgdown', 'pkgload', 'profvis', 'rcmdcheck', 'rlang', 'roxygen2', 'rversions', 'sessioninfo', 'testthat', 'urlchecker' are not available for package 'devtools'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/devtools'

The downloaded source packages are in

‘C:\Users\adlik\AppData\Local\Temp\RtmpWe4OTi\downloaded\_packages’

Loading required package: devtools

Installing package into ‘C:/Users/adlik/OneDrive/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

Warning: unable to access index for repository https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6:

cannot open URL 'https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6/PACKAGES'

Package which is only available in source form, and may need

compilation of C/C++/Fortran: ‘magrittr’

These will not be installed

Loading required package: magrittr

Installing package into ‘C:/Users/adlik/OneDrive/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

also installing the dependencies ‘fansi’, ‘utf8’, ‘cli’, ‘glue’, ‘lifecycle’, ‘magrittr’, ‘pillar’, ‘rlang’, ‘tibble’, ‘tidyselect’, ‘vctrs’

Warning: unable to access index for repository https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6:

cannot open URL 'https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6/PACKAGES'

Packages which are only available in source form, and may need

compilation of C/C++/Fortran: ‘fansi’ ‘utf8’ ‘cli’ ‘glue’ ‘magrittr’

‘rlang’ ‘tibble’ ‘tidyselect’ ‘vctrs’ ‘dplyr’

These will not be installed

installing the source packages ‘lifecycle’, ‘pillar’

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/lifecycle\_1.0.4.tar.gz'

Content type 'application/x-gzip' length 107656 bytes (105 KB)

downloaded 105 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/pillar\_1.9.0.tar.gz'

Content type 'application/x-gzip' length 444528 bytes (434 KB)

downloaded 434 KB

ERROR: dependencies 'cli', 'glue', 'rlang' are not available for package 'lifecycle'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/lifecycle'

ERROR: dependencies 'cli', 'fansi', 'glue', 'lifecycle', 'rlang', 'utf8', 'vctrs' are not available for package 'pillar'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/pillar'

The downloaded source packages are in

‘C:\Users\adlik\AppData\Local\Temp\RtmpWe4OTi\downloaded\_packages’

Loading required package: dplyr

Installing package into ‘C:/Users/adlik/OneDrive/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

Warning: unable to access index for repository https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6:

cannot open URL 'https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6/PACKAGES'

Loading required package: survey

This is mgcv 1.8-31. For overview type 'help("mgcv-package")'.

Installing package into ‘C:/Users/adlik/OneDrive/Documents/R/win-library/3.6’

(as ‘lib’ is unspecified)

Warning: dependency ‘rainbow’ is not available

also installing the dependencies ‘bitops’, ‘colorspace’, ‘utf8’, ‘RCurl’, ‘farver’, ‘munsell’, ‘fansi’, ‘magrittr’, ‘pillar’, ‘fds’, ‘deSolve’, ‘minqa’, ‘nloptr’, ‘Rcpp’, ‘RcppEigen’, ‘cli’, ‘glue’, ‘gtable’, ‘isoband’, ‘lifecycle’, ‘rlang’, ‘scales’, ‘tibble’, ‘vctrs’, ‘fda’, ‘mgcv’, ‘gamm4’, ‘lme4’, ‘RLRsim’, ‘grpreg’, ‘ggplot2’

Warning: unable to access index for repository https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6:

cannot open URL 'https://mirror.las.iastate.edu/CRAN/bin/windows/contrib/3.6/PACKAGES'

Packages which are only available in source form, and may need

compilation of C/C++/Fortran: ‘bitops’ ‘colorspace’ ‘utf8’ ‘RCurl’

‘farver’ ‘fansi’ ‘magrittr’ ‘deSolve’ ‘minqa’ ‘nloptr’ ‘Rcpp’

‘RcppEigen’ ‘cli’ ‘glue’ ‘isoband’ ‘rlang’ ‘scales’ ‘tibble’ ‘vctrs’

‘fda’ ‘mgcv’ ‘lme4’ ‘RLRsim’ ‘grpreg’

These will not be installed

installing the source packages ‘munsell’, ‘pillar’, ‘fds’, ‘gtable’, ‘lifecycle’, ‘gamm4’, ‘ggplot2’, ‘refund’

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/munsell\_0.5.1.tar.gz'

Content type 'application/x-gzip' length 182310 bytes (178 KB)

downloaded 178 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/pillar\_1.9.0.tar.gz'

Content type 'application/x-gzip' length 444528 bytes (434 KB)

downloaded 434 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/fds\_1.8.tar.gz'

Content type 'application/x-gzip' length 3860041 bytes (3.7 MB)

downloaded 3.7 MB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/gtable\_0.3.5.tar.gz'

Content type 'application/x-gzip' length 130227 bytes (127 KB)

downloaded 127 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/lifecycle\_1.0.4.tar.gz'

Content type 'application/x-gzip' length 107656 bytes (105 KB)

downloaded 105 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/gamm4\_0.2-6.tar.gz'

Content type 'application/x-gzip' length 14969 bytes (14 KB)

downloaded 14 KB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/ggplot2\_3.5.1.tar.gz'

Content type 'application/x-gzip' length 3604371 bytes (3.4 MB)

downloaded 3.4 MB

trying URL 'https://mirror.las.iastate.edu/CRAN/src/contrib/refund\_0.1-35.tar.gz'

Content type 'application/x-gzip' length 1753526 bytes (1.7 MB)

downloaded 1.7 MB

ERROR: dependency 'colorspace' is not available for package 'munsell'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/munsell'

ERROR: dependencies 'rainbow', 'RCurl' are not available for package 'fds'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/fds'

ERROR: dependencies 'cli', 'glue', 'rlang' are not available for package 'lifecycle'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/lifecycle'

ERROR: dependency 'lme4' is not available for package 'gamm4'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/gamm4'

ERROR: dependencies 'cli', 'fansi', 'glue', 'lifecycle', 'rlang', 'utf8', 'vctrs' are not available for package 'pillar'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/pillar'

ERROR: dependencies 'cli', 'glue', 'lifecycle', 'rlang' are not available for package 'gtable'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/gtable'

ERROR: dependencies 'cli', 'glue', 'gtable', 'isoband', 'lifecycle', 'rlang', 'scales', 'tibble', 'vctrs' are not available for package 'ggplot2'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/ggplot2'

ERROR: dependencies 'fda', 'gamm4', 'lme4', 'RLRsim', 'grpreg', 'ggplot2' are not available for package 'refund'

\* removing 'C:/Users/adlik/OneDrive/Documents/R/win-library/3.6/refund'

The downloaded source packages are in

‘C:\Users\adlik\AppData\Local\Temp\RtmpWe4OTi\downloaded\_packages’

Loading required package: refund

$tableone

[1] FALSE

$knitr

[1] FALSE

$kableExtra

[1] FALSE

$devtools

[1] FALSE

$magrittr

[1] FALSE

$dplyr

[1] FALSE

$survey

[1] FALSE

$mgcv

NULL

$refund

[1] FALSE

There were 50 or more warnings (use warnings() to see the first 50)

> rm(list=c("pckgs"))

>

> ## Install the rnhanesdata package and dependencies.

> ## This may take a few minutes because of the size of the data package.

> if(!require("rnhanesdata")){

+ install\_github("andrew-leroux/rnhanesdata")

+ require("rnhanesdata")

+ }

Loading required package: rnhanesdata

Error in install\_github("andrew-leroux/rnhanesdata") :

could not find function "install\_github"

In addition: Warning message:

In library(package, lib.loc = lib.loc, character.only = TRUE, logical.return = TRUE, :

there is no package called ‘rnhanesdata’

>

> make\_plots <- TRUE ## change to FALSE if you don't want to create the figures presented in the manuscript

> make\_tables <- TRUE ## change to FALSE if you don't want to create the tables presented in the manuscript

>

> ## change "make\_plot\_fpca\_vs\_svypca" below to TRUE if you want to plot the first 16 survey weighted principal components versus

> ## the unweighted functional principal components

> make\_plot\_fpca\_vs\_svypca <- FALSE

>

>

> dir\_supplemental <- "C:\Users\adlik\Documents\NHANESR20132014\dir\_supplemental\" ## directory where supplemental material folder is saved

Error: '\U' used without hex digits in character string starting ""C:\U"

> code\_path <- file.path(dir\_supplemental, "supplemental\_material", "code") ## file path where helper functions and code to create figures are located

> figure\_path <- file.path(dir\_supplemental, "supplemental\_material", "figures") ## file path where figures will be saved

> table\_path <- file.path(dir\_supplemental, "supplemental\_material", "tables") ## file path where tables will be saved

>

>

> ## Source a single helper function: calc\_weighted\_AUC().

> ## This function calculates survey weighted AUC given a set of labels and predictions

> source(file.path(code\_path,"helper\_fns.R"))

>

>

>

> #######################################

> ## ##

> ## Section 1a: load and merge data ##

> ## ##

> #######################################

>

> ## load the data

> data("PAXINTEN\_C");data("PAXINTEN\_D")

Warning message:

In data("PAXINTEN\_C") : data set ‘PAXINTEN\_C’ not found

Warning message:

In data("PAXINTEN\_D") : data set ‘PAXINTEN\_D’ not found

> data("Flags\_C");data("Flags\_D")

Warning message:

In data("Flags\_C") : data set ‘Flags\_C’ not found

Warning message:

In data("Flags\_D") : data set ‘Flags\_D’ not found

> data("Mortality\_2011\_C");data("Mortality\_2011\_D")

Warning message:

In data("Mortality\_2011\_C") : data set ‘Mortality\_2011\_C’ not found

Warning message:

In data("Mortality\_2011\_D") : data set ‘Mortality\_2011\_D’ not found

> data("Covariate\_C");data("Covariate\_D")

Warning message:

In data("Covariate\_C") : data set ‘Covariate\_C’ not found

Warning message:

In data("Covariate\_D") : data set ‘Covariate\_D’ not found

>

>

> ## re-code activity counts which are considered "non-wear" to be 0

> ## this doesn't impact much data, most estimated non-wear times correspond to 0 counts anyway

> PAXINTEN\_C[,paste0("MIN",1:1440)] <- PAXINTEN\_C[,paste0("MIN",1:1440)]\*Flags\_C[,paste0("MIN",1:1440)]

Error: object 'PAXINTEN\_C' not found

> PAXINTEN\_D[,paste0("MIN",1:1440)] <- PAXINTEN\_D[,paste0("MIN",1:1440)]\*Flags\_D[,paste0("MIN",1:1440)]

Error: object 'PAXINTEN\_D' not found

>

>

>

> ## Merge covariate, mortality, and accelerometry data

> ## note that both PAXINTEN\_\* and Covariate\_\* have a column

> ## called "SDDSRVYR" indicating which NHANES wave the data is associated with.

> ## To avoid duplicating this column in the merged data, we add this variable to the "by"

> ## argument in left\_join()

> AllAct\_C <- left\_join(PAXINTEN\_C, Mortality\_2011\_C, by = "SEQN") %>%

+ left\_join(Covariate\_C, by=c("SEQN", "SDDSRVYR"))

Error in left\_join(PAXINTEN\_C, Mortality\_2011\_C, by = "SEQN") %>% left\_join(Covariate\_C, :

could not find function "%>%"

> AllAct\_D <- left\_join(PAXINTEN\_D, Mortality\_2011\_D, by = "SEQN") %>%

+ left\_join(Covariate\_D, by=c("SEQN", "SDDSRVYR"))

Error in left\_join(PAXINTEN\_D, Mortality\_2011\_D, by = "SEQN") %>% left\_join(Covariate\_D, :

could not find function "%>%"

>

> AllFlags\_C <- left\_join(Flags\_C, Mortality\_2011\_C, by = "SEQN") %>%

+ left\_join(Covariate\_C, by=c("SEQN", "SDDSRVYR"))

Error in left\_join(Flags\_C, Mortality\_2011\_C, by = "SEQN") %>% left\_join(Covariate\_C, :

could not find function "%>%"

> AllFlags\_D <- left\_join(Flags\_D, Mortality\_2011\_D, by = "SEQN") %>%

+ left\_join(Covariate\_D, by=c("SEQN", "SDDSRVYR"))

Error in left\_join(Flags\_D, Mortality\_2011\_D, by = "SEQN") %>% left\_join(Covariate\_D, :

could not find function "%>%"

>

> ## clean up the workspace for memory purposes

> rm(list=c(paste0(c("PAXINTEN\_", "Covariate\_","Mortality\_2011\_","Flags\_"),rep(LETTERS[3:4],each=4))))

Warning messages:

1: In rm(list = c(paste0(c("PAXINTEN\_", "Covariate\_", "Mortality\_2011\_", :

object 'PAXINTEN\_C' not found

2: In rm(list = c(paste0(c("PAXINTEN\_", "Covariate\_", "Mortality\_2011\_", :

object 'Covariate\_C' not found

3: In rm(list = c(paste0(c("PAXINTEN\_", "Covariate\_", "Mortality\_2011\_", :

object 'Mortality\_2011\_C' not found

4: In rm(list = c(paste0(c("PAXINTEN\_", "Covariate\_", "Mortality\_2011\_", :

object 'Flags\_C' not found

5: In rm(list = c(paste0(c("PAXINTEN\_", "Covariate\_", "Mortality\_2011\_", :

object 'PAXINTEN\_D' not found

6: In rm(list = c(paste0(c("PAXINTEN\_", "Covariate\_", "Mortality\_2011\_", :

object 'Covariate\_D' not found

7: In rm(list = c(paste0(c("PAXINTEN\_", "Covariate\_", "Mortality\_2011\_", :

object 'Mortality\_2011\_D' not found

8: In rm(list = c(paste0(c("PAXINTEN\_", "Covariate\_", "Mortality\_2011\_", :

object 'Flags\_D' not found

>

> ## combine data for the two waves

> AllAct <- rbind.data.frame(AllAct\_C,AllAct\_D)

Error in rbind.data.frame(AllAct\_C, AllAct\_D) :

object 'AllAct\_C' not found

> AllFlags <- rbind.data.frame(AllFlags\_C,AllFlags\_D)

Error in rbind.data.frame(AllFlags\_C, AllFlags\_D) :

object 'AllFlags\_C' not found

>

>

> ## clean up the workspace again

> rm(list=c("AllAct\_C","AllAct\_D","AllFlags\_C","AllFlags\_D"))

Warning messages:

1: In rm(list = c("AllAct\_C", "AllAct\_D", "AllFlags\_C", "AllFlags\_D")) :

object 'AllAct\_C' not found

2: In rm(list = c("AllAct\_C", "AllAct\_D", "AllFlags\_C", "AllFlags\_D")) :

object 'AllAct\_D' not found

3: In rm(list = c("AllAct\_C", "AllAct\_D", "AllFlags\_C", "AllFlags\_D")) :

object 'AllFlags\_C' not found

4: In rm(list = c("AllAct\_C", "AllAct\_D", "AllFlags\_C", "AllFlags\_D")) :

object 'AllFlags\_D' not found

>

>

>

>

> ##############################################################################

> ## ##

> ## Section 1b: create new variables/relevel factor variables for analyses ##

> ## ##

> ##############################################################################

>

> ## Code year 5 mortality, NAs for individuals with follow up less than 5 years and alive

> AllAct$yr5\_mort <- AllFlags$yr5\_mort <- as.integer(ifelse(AllAct$permth\_exm/12 <= 5 & AllAct$mortstat == 1, 1,

+ ifelse(AllAct$permth\_exm/12 < 5 & AllAct$mortstat == 0, NA, 0))

+ )

Error in ifelse(AllAct$permth\_exm/12 <= 5 & AllAct$mortstat == 1, 1, ifelse(AllAct$permth\_exm/12 < :

object 'AllAct' not found

>

> ## Create Age in years using the age at examination (i.e. when participants wore the device)

> AllAct$Age <- AllFlags$Age <- AllAct$RIDAGEEX/12

Error: object 'AllAct' not found

>

> ## Re-level comorbidities to assign refused/don't know as not having the condition

> ## Note that in practice this does not affect many individuals, but it is an assumption we're making.

> levels(AllAct$CHD) <- levels(AllFlags$CHD) <- list("No" = c("No","Refused","Don't know"), "Yes" = c("Yes"))

Error in levels(AllFlags$CHD) <- list(No = c("No", "Refused", "Don't know"), :

object 'AllFlags' not found

> levels(AllAct$CHF) <- levels(AllFlags$CHF) <- list("No" = c("No","Refused","Don't know"), "Yes" = c("Yes"))

Error in levels(AllFlags$CHF) <- list(No = c("No", "Refused", "Don't know"), :

object 'AllFlags' not found

> levels(AllAct$Stroke) <- levels(AllFlags$Stroke) <- list("No" = c("No","Refused","Don't know"), "Yes" = c("Yes"))

Error in levels(AllFlags$Stroke) <- list(No = c("No", "Refused", "Don't know"), :

object 'AllFlags' not found

> levels(AllAct$Cancer) <- levels(AllFlags$Cancer) <- list("No" = c("No","Refused","Don't know"), "Yes" = c("Yes"))

Error in levels(AllFlags$Cancer) <- list(No = c("No", "Refused", "Don't know"), :

object 'AllFlags' not found

> levels(AllAct$Diabetes) <- levels(AllFlags$Diabetes) <- list("No" = c("No","Borderline", "Refused","Don't know"), "Yes" = c("Yes"))

Error in levels(AllFlags$Diabetes) <- list(No = c("No", "Borderline", :

object 'AllFlags' not found

>

>

> ## Re-level education to have 3 levels and categorize don't know/refused to be missing

> levels(AllAct$EducationAdult) <- levels(AllFlags$EducationAdult) <- list("Less than high school" = c("Less than 9th grade", "9-11th grade"),

+ "High school" = c("High school grad/GED or equivalent"),

+ "More than high school" = c("Some College or AA degree", "College graduate or above"))

Error in levels(AllFlags$EducationAdult) <- list(`Less than high school` = c("Less than 9th grade", :

object 'AllFlags' not found

>

> ## Re-level alcohol consumption to include a level for "missing"

> levels(AllAct$DrinkStatus) <- levels(AllFlags$DrinkStatus) <- c(levels(AllAct$DrinkStatus), "Missing alcohol")

Error in levels(AllAct$DrinkStatus) : object 'AllAct' not found

> AllAct$DrinkStatus[is.na(AllAct$DrinkStatus)] <- AllFlags$DrinkStatus[is.na(AllAct$DrinkStatus)] <- "Missing alcohol"

Error in AllFlags$DrinkStatus[is.na(AllAct$DrinkStatus)] <- "Missing alcohol" :

object 'AllFlags' not found

>

>

>

> ## Re-order columns so that activity and wear/non-wear flags are the last 1440 columns of our two

> ## data matrices. This is a personal preference and is absolutely not necessary.

> act\_cols <- which(colnames(AllAct) %in% paste0("MIN",1:1440))

Error in is.data.frame(x) : object 'AllAct' not found

> oth\_cols <- which(!colnames(AllAct) %in% paste0("MIN",1:1440))

Error in is.data.frame(x) : object 'AllAct' not found

> AllAct <- AllAct[,c(oth\_cols,act\_cols)]

Error: object 'AllAct' not found

> AllFlags <- AllFlags[,c(oth\_cols,act\_cols)]

Error: object 'AllFlags' not found

> rm(list=c("act\_cols","oth\_cols"))

Warning messages:

1: In rm(list = c("act\_cols", "oth\_cols")) : object 'act\_cols' not found

2: In rm(list = c("act\_cols", "oth\_cols")) : object 'oth\_cols' not found

>

>

>

>

>

>

>

> ###########################################################

> ## ##

> ## Section 2: Calcualte common accelerometery features ##

> ## ##

> ###########################################################

>

> ## Assign just the activity and wear/non-wear flag data to matrices.

> ## This makes computing the features faster but is technically required.

> act\_mat <- as.matrix(AllAct[,paste0("MIN",1:1440)])

Error in as.matrix(AllAct[, paste0("MIN", 1:1440)]) :

object 'AllAct' not found

> flag\_mat <- as.matrix(AllFlags[,paste0("MIN",1:1440)])

Error in as.matrix(AllFlags[, paste0("MIN", 1:1440)]) :

object 'AllFlags' not found

>

> ## replace NAs with 0s

> ## As described in the manuscript, this only affects 501 minutes for 1 day, for one subject

> act\_mat[is.na(act\_mat)] <- 0

Error in act\_mat[is.na(act\_mat)] <- 0 : object 'act\_mat' not found

> flag\_mat[is.na(flag\_mat)] <- 0

Error in flag\_mat[is.na(flag\_mat)] <- 0 : object 'flag\_mat' not found

>

>

> AllAct$TAC <- AllFlags$TAC <- rowSums(act\_mat)

Error in is.data.frame(x) : object 'act\_mat' not found

> AllAct$TLAC <- AllFlags$TLAC <- rowSums(log(1+act\_mat))

Error in is.data.frame(x) : object 'act\_mat' not found

> AllAct$WT <- AllFlags$WT <- rowSums(flag\_mat)

Error in is.data.frame(x) : object 'flag\_mat' not found

> AllAct$ST <- AllFlags$ST <- rowSums(act\_mat < 100)

Error in is.data.frame(x) : object 'act\_mat' not found

> AllAct$MVPA <- AllFlags$MVPA <- rowSums(act\_mat >= 2020)

Error in is.data.frame(x) : object 'act\_mat' not found

>

> ## calculate fragmentation measures

> bout\_mat <- apply(act\_mat >= 100, 1, function(x){

+ mat <- rle(x)

+ sed <- mat$lengths[which(mat$values == FALSE)]

+ act <- mat$length[mat$values == TRUE]

+

+ sed <- ifelse(length(sed) == 0, NA, mean(sed))

+ act <- ifelse(length(act) == 0, NA, mean(act))

+ c(sed,act)

+ })

Error in apply(act\_mat >= 100, 1, function(x) { :

object 'act\_mat' not found

>

> AllAct$SBout <- AllFlags$SBout <- bout\_mat[1,]

Error: object 'bout\_mat' not found

> AllAct$ABout <- AllFlags$ABout <- bout\_mat[2,]

Error: object 'bout\_mat' not found

> AllAct$SATP <- AllFlags$SATP <- 1/AllAct$SBout

Error: object 'AllAct' not found

> AllAct$ASTP <- AllFlags$ASTP <- 1/AllAct$ABout

Error: object 'AllAct' not found

> rm(list=c("act\_mat","flag\_mat","bout\_mat"))

Warning messages:

1: In rm(list = c("act\_mat", "flag\_mat", "bout\_mat")) :

object 'act\_mat' not found

2: In rm(list = c("act\_mat", "flag\_mat", "bout\_mat")) :

object 'flag\_mat' not found

3: In rm(list = c("act\_mat", "flag\_mat", "bout\_mat")) :

object 'bout\_mat' not found

>

>

>

> ###########################################

> ## ##

> ## Section 3: Apply exclusion criteria ##

> ## ##

> ###########################################

>

>

> ## make dataframe with one row per individual to create table 1.

> ## Remove columns associated with activity to avoid any confusion.

> table\_dat <- AllAct[!duplicated(AllAct$SEQN),-which(colnames(AllAct) %in% c(paste0("MIN",1:1440),

+ "TAC","TLAC","WT","ST","MVPA",

+ "SBout","ABout","SATP","ASTP"))]

Error: object 'AllAct' not found

>

> ## subset based on our age inclusion/exclusion criteria

> ## note that individuals age 85 and over are coded as NA

> table\_dat <- subset(table\_dat, !(Age < 50 | is.na(Age)))

>

> ## get the SEQN (id variable) associated with individuals with fewer than 3 days accelerometer wear time

> ## with at least 10 hours OR had their data quality/device calibration flagged by NHANES

> keep\_inx <- exclude\_accel(AllAct, AllFlags)

Error in exclude\_accel(AllAct, AllFlags) :

could not find function "exclude\_accel"

> Act\_Analysis <- AllAct[keep\_inx,]

Error: object 'AllAct' not found

> Flags\_Analysis <- AllFlags[keep\_inx,]

Error: object 'AllFlags' not found

> nms\_rm <- unique(c(Act\_Analysis$SEQN[-which(Act\_Analysis$SEQN %in% names(table(Act\_Analysis$SEQN))[table(Act\_Analysis$SEQN)>=3])],

+ setdiff(AllAct$SEQN, Act\_Analysis$SEQN))

+ )

Error in as.vector(x) : object 'AllAct' not found

> rm(list=c("keep\_inx"))

Warning message:

In rm(list = c("keep\_inx")) : object 'keep\_inx' not found

>

>

> ## Additional inclusion/exclusion criteria.

> ## Aside from mortality or accelerometer weartime, the only missingness is in

> ## Education (6) and BMI (35).

> criteria\_vec <- c("(is.na(table\_dat$BMI\_cat))", # missing BMI

+ "(is.na(table\_dat$EducationAdult))", # missing education

+ "(table\_dat$SEQN %in% nms\_rm)", # too few "good" days of accelerometery data

+ "((!table\_dat$eligstat %in% 1) | is.na(table\_dat$mortstat) | is.na(table\_dat$permth\_exm) | table\_dat$ucod\_leading %in% \"004\")", # missing mortality data, or accidental death

+ "(table\_dat$mortstat == 0 & table\_dat$permth\_exm/12 < 5)" # less than 5 years of follow up with no mortality

+ )

>

> ## create matrix of pairwise missing data based on our exclusion criterial

> tab\_miss <- matrix(NA, ncol=length(criteria\_vec), nrow=length(criteria\_vec))

> for(i in seq\_along(criteria\_vec)){

+ for(j in seq\_along(criteria\_vec)){

+ eval(parse(text=paste0("miss\_cur <- which(", criteria\_vec[i], "&", criteria\_vec[j],")")))

+ tab\_miss[i,j] <- length(miss\_cur)

+ rm(list=c("miss\_cur"))

+ }

+ }

Error in table\_dat$SEQN %in% nms\_rm : object 'nms\_rm' not found

> rownames(tab\_miss) <- colnames(tab\_miss) <- c("BMI","Education","Bad Accel Data","Mortality","Follow-up")

> rm(list=c("i","j"))

> ## view missing data pattern

> tab\_miss

BMI Education Bad Accel Data Mortality Follow-up

BMI 35 0 NA NA NA

Education NA NA NA NA NA

Bad Accel Data NA NA NA NA NA

Mortality NA NA NA NA NA

Follow-up NA NA NA NA NA

>

>

> ## add in column indicating exclusion:

> ## Exclude = 1 indicates an individual does not meet our inclusion criteria

> ## Exclude = 0 indicates an individual does meet our inclusion criteria

> eval(parse(text=paste0("table\_dat$Exclude <- as.integer(", paste0(criteria\_vec,collapse="|"), ")")))

Error in table\_dat$SEQN %in% nms\_rm : object 'nms\_rm' not found

>

>

> ## Create our dataset for analysis with one row per subject

> ## containing only those subjects who meet our inclusion criteria.

> data\_analysis <- subset(table\_dat, Exclude == 0)

> ## get adjusted survey weights using the reweight\_accel function

> data\_analysis <- reweight\_accel(data\_analysis)

Error in reweight\_accel(data\_analysis) :

could not find function "reweight\_accel"

>

> ## Get activity/flag data for only those included participants AND who have 3 good days of data.

> ## Since we've already removed the "bad" days from Act\_Analysis and Act\_Flags,

> ## we need only subset based on subject ID now.

> Act\_Analysis <- subset(Act\_Analysis, SEQN %in% data\_analysis$SEQN)

> Flags\_Analysis <- subset(Flags\_Analysis, SEQN %in% data\_analysis$SEQN)

>

> ## calculate subject specific averages of the accelerometry features

> ## using only the "good" days of data

> act\_var\_nms <- c("TAC","TLAC","WT","ST","MVPA","SATP","ASTP")

> for(i in act\_var\_nms){

+ data\_analysis[[i]] <- vapply(data\_analysis$SEQN, function(x) mean(Act\_Analysis[[i]][Act\_Analysis$SEQN==x],na.rm=TRUE), numeric(1))

+ }

>

> ## verify there's no missingness in the rest of our predictors of interest

> vars\_interest <- c("Age", "Gender", "Race", "EducationAdult", "SmokeCigs", "DrinkStatus", "BMI\_cat",

+ "Diabetes","CHF", "CHD", "Stroke",

+ "Cancer", "MobilityProblem",

+ "permth\_exm")

> summary(data\_analysis[,c(vars\_interest,act\_var\_nms,"mortstat")])

Age Gender Race

Min. :50.00 Male :1611 White :1862

1st Qu.:58.08 Female:1587 Mexican American: 572

Median :65.50 Other Hispanic : 61

Mean :65.97 Black : 601

3rd Qu.:73.56 Other : 102

Max. :84.92

EducationAdult SmokeCigs DrinkStatus

Less than high school:1021 Never :1430 Moderate Drinker:1512

High school : 792 Former :1229 Non-Drinker :1362

More than high school:1385 Current: 539 Heavy Drinker : 188

Missing alcohol : 136

BMI\_cat Diabetes CHF CHD Stroke Cancer

Normal : 827 No :2628 No :3009 No :2933 No :3006 No :2695

Underweight: 30 Yes: 570 Yes: 189 Yes: 265 Yes: 192 Yes: 503

Overweight :1218

Obese :1123

MobilityProblem permth\_exm TAC TLAC

No Difficulty :2163 Min. : 1.00 Min. : 8932 Min. : 429.9

Any Difficulty:1035 1st Qu.: 66.00 1st Qu.:124117 1st Qu.:2235.9

Median : 80.00 Median :190580 Median :2740.6

Mean : 77.27 Mean :206781 Mean :2741.1

3rd Qu.: 92.00 3rd Qu.:266993 3rd Qu.:3225.9

Max. :107.00 Max. :912076 Max. :5655.5

WT ST MVPA SATP

Min. : 615.3 Min. : 648.6 Min. : 0.000 Min. :0.007614

1st Qu.: 792.5 1st Qu.:1043.0 1st Qu.: 2.333 1st Qu.:0.066523

Median : 859.7 Median :1115.5 Median : 7.143 Median :0.080574

Mean : 878.2 Mean :1113.3 Mean : 13.534 Mean :0.082059

3rd Qu.: 930.1 3rd Qu.:1190.0 3rd Qu.: 18.857 3rd Qu.:0.095511

Max. :1440.0 Max. :1425.0 Max. :152.333 Max. :0.199420

ASTP mortstat

Min. :0.05228 Min. :0.0000

1st Qu.:0.24118 1st Qu.:0.0000

Median :0.29057 Median :0.0000

Mean :0.30182 Mean :0.1563

3rd Qu.:0.34816 3rd Qu.:0.0000

Max. :0.74495 Max. :1.0000

>

> write.csv(data\_analysis, file = "NHANES\_analysis\_data3.csv", row.names = TRUE)

>

>

> ## clean up the workspace

> rm(list=c("AllAct","AllFlags","i","criteria\_vec","nms\_rm","tab\_miss"))

Warning messages:

1: In rm(list = c("AllAct", "AllFlags", "i", "criteria\_vec", "nms\_rm", :

object 'AllAct' not found

2: In rm(list = c("AllAct", "AllFlags", "i", "criteria\_vec", "nms\_rm", :

object 'AllFlags' not found

3: In rm(list = c("AllAct", "AllFlags", "i", "criteria\_vec", "nms\_rm", :

object 'nms\_rm' not found

> gc()

used (Mb) gc trigger (Mb) max used (Mb)

Ncells 1482857 79.2 2315417 123.7 2315417 123.7

Vcells 63398625 483.7 108870872 830.7 108870644 830.7

>

>

>

> #####################################################

> # Table 1: Compare included vs Excluded individuals #

> #####################################################

> if(make\_tables){

+ source(file.path(code\_path, "create\_table1.R"))

+ }

Error in reweight\_accel(table\_dat) :

could not find function "reweight\_accel"

>

>

>

>

>

> ###################################

> ## ##

> ## Section 4: Data Analysis ##

> ## ##

> ###################################

>

>

> #############################

> ## ##

> ## Section 4.a: PCA/fPCA ##

> ## ##

> #############################

>

> ## get activity in matrix format, log transform, add 1 because of 0 counts

> ## There is one individual with 501 minutes recorded as NA. These missing data occur on the last day they wore the device

> ## for the last 501 minutes of the day. We impute these missing data with 0.

> Act <- as.matrix(log(1+Act\_Analysis[,paste0("MIN",1:1440)]))

> Act[is.na(Act)] <- 0

>

> start <- Sys.time()

> fpca\_fit <- fpca.face(Act,knots=50)

Error in fpca.face(Act, knots = 50) : could not find function "fpca.face"

> (end <- Sys.time()-start)

Time difference of 0.0312531 secs

> rm(list=c("start","end"))

>

> ## exmine the proportion of variability explained by the first 6 principal components

> c(cumsum(fpca\_fit$evalues)/sum(fpca\_fit$evalues))[1:6]

Error: object 'fpca\_fit' not found

>

>

> ##calculate the mean and sd of the first 6 PC scores for all participants

> for(k in 1:6){

+ data\_analysis[[paste0("mi",k)]] <- vapply(data\_analysis$SEQN, function(x) mean(fpca\_fit$scores[Act\_Analysis$SEQN==x,k]), numeric(1))

+ data\_analysis[[paste0("si",k)]] <- vapply(data\_analysis$SEQN, function(x) sd(fpca\_fit$scores[Act\_Analysis$SEQN==x,k]), numeric(1))

+ }

Error in mean(fpca\_fit$scores[Act\_Analysis$SEQN == x, k]) :

object 'fpca\_fit' not found

>

>

> ##########################################

> ## Figure: first 6 principal components ##

> ##########################################

>

> if(make\_plots){

+ source(file.path(code\_path,"create\_principal\_components\_plot.R"))

+ if(make\_plot\_fpca\_vs\_svypca){

+ source(file.path(code\_path,"create\_svyprincipal\_components\_plot.R"))

+ }

+ }

Error in apply(fpca\_fit$scores, 2, sd) : object 'fpca\_fit' not found

> rm(list=c("make\_plot\_fpca\_vs\_svypca"))

>

>

>

> ############################################################################

> ## ##

> ## Section 4.b: Backward selection to identify features of first 6 fPCs ##

> ## associated with 5-year mortality ##

> ## ##

> ############################################################################

>

> ## Create a svydesign() object for

> ## estimating complex survey generalized linear models.

> ## Here we use the adjusted (re-weighted) 4-year normalized survey weights.

> data\_analysis\_svy <- svydesign(id= ~SDMVPSU, strata = ~SDMVSTRA,

+ weights = ~wtmec4yr\_adj\_norm, data = data\_analysis, nest = TRUE)

Error in svydesign(id = ~SDMVPSU, strata = ~SDMVSTRA, weights = ~wtmec4yr\_adj\_norm, :

could not find function "svydesign"

>

> ind\_vars <- c("mi1", "mi2","mi3", "mi4","mi5","mi6",

+ "si1","si2","si3","si4","si5", "si6")

> inc\_vars <- c("Age", "SmokeCigs", "DrinkStatus", "BMI\_cat",

+ "Diabetes","CHF", "CHD", "Stroke",

+ "Cancer", "MobilityProblem")

> exc\_vars <- ind\_vars

> aic\_vec <- var\_vec <- model\_vec <- rep(NA, length(ind\_vars))

> for(i in 1:length(ind\_vars)){

+ aic\_ij <- rep(NA,length(exc\_vars))

+

+ for(k in 1:length(exc\_vars)){

+ form <- paste0(c(inc\_vars, exc\_vars[-k]), collapse="+")

+ fit\_tmp <- svyglm(as.formula(paste("yr5\_mort ~", form)), design=data\_analysis\_svy,family=quasibinomial())

+

+ aic\_ij[k] <- AIC(fit\_tmp)[2]

+

+ rm(list=c("fit\_tmp","form"))

+ }

+

+ k\_cur <- which(aic\_ij == min(aic\_ij))

+ model\_vec[i] <- paste0(c(inc\_vars, exc\_vars[-k\_cur]), collapse="+")

+ exc\_vars <- exc\_vars[-k\_cur]

+ aic\_vec[i] <- aic\_ij[k\_cur]

+ rm(list=c("k\_cur","aic\_ij","k"))

+ }

Error in svyglm(as.formula(paste("yr5\_mort ~", form)), design = data\_analysis\_svy, :

could not find function "svyglm"

> ## get the final model as the first model where AIC increases after removing a variable

> ## identified mi1, si1, si5, si6

> (backward\_model <- model\_vec[which(diff(aic\_vec) > 0) + 1][1])

[1] NA

>

> fit\_logistic\_pca <- svyglm(as.formula(paste("yr5\_mort ~", backward\_model)), design=data\_analysis\_svy,

+ family=quasibinomial())

Error in svyglm(as.formula(paste("yr5\_mort ~", backward\_model)), design = data\_analysis\_svy, :

could not find function "svyglm"

> summary(fit\_logistic\_pca)

Error in summary(fit\_logistic\_pca) : object 'fit\_logistic\_pca' not found

> rm(list=c("aic\_vec","i","model\_vec","exc\_vars","ind\_vars","var\_vec",

+ "inc\_vars","data\_analysis\_svy","backward\_model","fit\_logistic\_pca"))

Warning messages:

1: In rm(list = c("aic\_vec", "i", "model\_vec", "exc\_vars", "ind\_vars", :

object 'data\_analysis\_svy' not found

2: In rm(list = c("aic\_vec", "i", "model\_vec", "exc\_vars", "ind\_vars", :

object 'fit\_logistic\_pca' not found

>

>

>

>

> ## Having identified candidate predictors, we now attempt to find surrogate measures for these PC

> ## features on the original data.

> ## To qualify as a surrogate, the feature calculated on the raw data should be highly correlated with the PC

> ## feature it is associated with (~|0.75|+).

> ## The procedure for finding these surrogate measures involved a fair amount of "guess and check" based on the shape

> ## of the principal components. The "guessing and checking" is not shown here in the interest of conciseness. We discuss this idea in a bit more detail in the manuscript.

> ##

> ## Our search here is by no means exhaustive, and we focused primarily on "average" activity during various

> ## periods of the day indicated by the shape of each principal component.

>

> #################################

> ## surrogate for m\_{i1}/s\_{i1} ##

> #################################

>

> ## Note that the shape of PC1 is predominantly a shift in average activity profile

> ## so we use TLAC as our surrogate measure for m\_{i1}, and the day-to-day standard deviation

> ## of TLAC as our surrogate for s\_{i1}

> data\_analysis$sPC1 <- tapply(Act\_Analysis$TLAC, Act\_Analysis$SEQN, sd)

>

> ## compare association between m\_{i1}/s\_{i1} with our surrogate measures

> cor(data\_analysis$TLAC,data\_analysis$mi1)

Error in cor(data\_analysis$TLAC, data\_analysis$mi1) :

supply both 'x' and 'y' or a matrix-like 'x'

> cor(data\_analysis$sPC1,data\_analysis$si1)

Error in cor(data\_analysis$sPC1, data\_analysis$si1) :

supply both 'x' and 'y' or a matrix-like 'x'

>

>

>

> ##########################

> ## surrogate for s\_{i5} ##

> ##########################

> t5a <- (11\*60+1):(15\*60)

> t5b <- (16\*60+1):(19\*60)

> s5a <- (rowMeans(Act[,t5a]) - rowMeans(Act[,t5b]))

> data\_analysis$sPC5 <- tapply(s5a, Act\_Analysis$SEQN, sd)

> cor(data\_analysis$sPC5, data\_analysis$si5)

Error in cor(data\_analysis$sPC5, data\_analysis$si5) :

supply both 'x' and 'y' or a matrix-like 'x'

> rm(list=c("t5a","t5b","s5a"))

>

>

>

> ##########################

> ## surrogate for s\_{i6} ##

> ##########################

> t6a <- c((8\*60+1):(10\*60),

+ (15\*60+1):(17\*60),

+ (22\*60+1):(24\*60))

> t6b <-c((5\*60+1):(7\*60),

+ (11\*60+1):(13\*60),

+ (18\*60+1):(20\*60))

> s6a <- rowMeans(Act[,t6a])-rowMeans(Act[,t6b])

> data\_analysis$sPC6 <- tapply(s6a, Act\_Analysis$SEQN, sd)

> cor(data\_analysis$sPC6,data\_analysis$si6)

Error in cor(data\_analysis$sPC6, data\_analysis$si6) :

supply both 'x' and 'y' or a matrix-like 'x'

> rm(list=c("t6a","t6b","s6a"))

>

>

>

>

>

>

> ##########################################################

> ## ##

> ## Section 4.c: Scalar on function regression (SoFR) ##

> ## ##

> ##########################################################

>

> ## get individual average activity and day-to-day variability profiles

> ## using log transformed activity smoothed via fPCA

> avg\_profiles <- matrix(NA, nrow=nrow(data\_analysis),ncol=ncol(Act))

> uid <- unique(data\_analysis$SEQN)

> for(i in 1:length(uid)){

+ avg\_profiles[i,] <- colMeans(fpca\_fit$Yhat[Act\_Analysis$SEQN == uid[i],])

+ }

Error in is.data.frame(x) : object 'fpca\_fit' not found

> ## center the functions so that our functional regression is identifiable

> ## add the centered functions to our dataframe using the I() function which retains matrix structure

> ## within the dataframe

> data\_analysis$X\_cn <- I(avg\_profiles - rep(1, nrow(avg\_profiles)) %\*% t(colMeans(avg\_profiles)))

> rm(list=c("uid","i"))

>

>

> ## fit the functional linear model using a cyclic p-spline basis

> ## Note that although this model accounts for survey weights, it ignores the clustering

> ## structure.

> fit\_SoFR <- pfr(yr5\_mort ~ Age + BMI\_cat + SmokeCigs + DrinkStatus + Race + EducationAdult+

+ CHD + Diabetes + CHF + Stroke +MobilityProblem+ Cancer +

+ lf(X\_cn, k=30, bs="cp"),

+ data=data\_analysis, family=quasibinomial(), weights=data\_analysis$wtmec4yr\_adj\_norm)

Error in pfr(yr5\_mort ~ Age + BMI\_cat + SmokeCigs + DrinkStatus + Race + :

could not find function "pfr"

>

> ## remove all variable names that got assigned to environment, possibly a bug with pfr?

> rm(list=ls()[ls()%in%colnames(data\_analysis)])

>

>

> ####################################

> ## Figure: Functional coefficient ##

> ####################################

> if(make\_plots){

+ source(file.path(code\_path,"create\_functional\_coefficient\_plot.R"))

+ }

Error in predict(fit\_SoFR, newdata = subset(data\_analysis, SEQN %in% subjs), :

object 'fit\_SoFR' not found

>

> ## remove the centered data matrix from our data frame

> ## this is only done to increase computational speed of the cross-validation below

> data\_analysis <- data\_analysis[,-which(colnames(data\_analysis) %in% c("X\_cn"))]

>

> rm(list=c("fpca\_fit","avg\_profiles","fit\_SoFR"))

Warning messages:

1: In rm(list = c("fpca\_fit", "avg\_profiles", "fit\_SoFR")) :

object 'fpca\_fit' not found

2: In rm(list = c("fpca\_fit", "avg\_profiles", "fit\_SoFR")) :

object 'fit\_SoFR' not found

>

>

>

>

>

>

>

>

> #################################################################################

> ## ##

> ## Section 4.d: Forward selection to identify predictive value of variables ##

> ## ##

> #################################################################################

>

>

>

>

> ## In this section we perform forward selection using the adjusted (re-weighted) survey weights,

> ## unadjusted weights, and no survey weights. We identify variables and output associated with these models

> ## using the naming convention: \_adj, \_unadj, \_wnwgt

> ## to denote adjusted, unadjusted, and unweighted results, respectively.

>

>

>

> ## variables to consider as linear predictors of 5-year mortality

> ind\_vars <- c("Age", "Gender", "Race", "EducationAdult", "SmokeCigs", "DrinkStatus", "BMI\_cat",

+ "Diabetes","CHF", "CHD", "Stroke", "Cancer", "MobilityProblem",

+ "sPC1", #"sPC5",

+ "sPC6", "ST", "WT", "MVPA","TAC", "TLAC", "SATP", "ASTP")

>

> ## examine distrubtion of predictor variables

> summary(data\_analysis[,ind\_vars])

Age Gender Race

Min. :50.00 Male :1611 White :1862

1st Qu.:58.08 Female:1587 Mexican American: 572

Median :65.50 Other Hispanic : 61

Mean :65.97 Black : 601

3rd Qu.:73.56 Other : 102

Max. :84.92

EducationAdult SmokeCigs DrinkStatus

Less than high school:1021 Never :1430 Moderate Drinker:1512

High school : 792 Former :1229 Non-Drinker :1362

More than high school:1385 Current: 539 Heavy Drinker : 188

Missing alcohol : 136

BMI\_cat Diabetes CHF CHD Stroke Cancer

Normal : 827 No :2628 No :3009 No :2933 No :3006 No :2695

Underweight: 30 Yes: 570 Yes: 189 Yes: 265 Yes: 192 Yes: 503

Overweight :1218

Obese :1123

MobilityProblem sPC1 sPC6 ST

No Difficulty :2163 Min. : 28.85 Min. :0.06039 Min. : 648.6

Any Difficulty:1035 1st Qu.: 333.52 1st Qu.:0.49361 1st Qu.:1043.0

Median : 469.30 Median :0.65368 Median :1115.5

Mean : 513.64 Mean :0.68558 Mean :1113.3

3rd Qu.: 633.38 3rd Qu.:0.84752 3rd Qu.:1190.0

Max. :2105.08 Max. :2.03188 Max. :1425.0

WT MVPA TAC TLAC

Min. : 615.3 Min. : 0.000 Min. : 8932 Min. : 429.9

1st Qu.: 792.5 1st Qu.: 2.333 1st Qu.:124117 1st Qu.:2235.9

Median : 859.7 Median : 7.143 Median :190580 Median :2740.6

Mean : 878.2 Mean : 13.534 Mean :206781 Mean :2741.1

3rd Qu.: 930.1 3rd Qu.: 18.857 3rd Qu.:266993 3rd Qu.:3225.9

Max. :1440.0 Max. :152.333 Max. :912076 Max. :5655.5

SATP ASTP

Min. :0.007614 Min. :0.05228

1st Qu.:0.066523 1st Qu.:0.24118

Median :0.080574 Median :0.29057

Mean :0.082059 Mean :0.30182

3rd Qu.:0.095511 3rd Qu.:0.34816

Max. :0.199420 Max. :0.74495

>

> ## standardize continuous predictors other than age

> ## so coefficients are interpretable as

> ## "a one standard deviation change in ... is associated with ... (in/de)crease in log odds of 5-year mortality"

> vars\_std <- c("sPC1", "sPC5", "sPC6", "ST", "WT", "MVPA","TAC", "TLAC", "SATP", "ASTP")

> for(i in vars\_std){

+ data\_analysis[[i]] <- data\_analysis[[i]]/sd(data\_analysis[[i]])

+ }

> rm(list=c("vars\_std","i"))

>

> ## Create svydesign() objects associated with the adjusted and unadjusted

> ## survey weights. These are the basis for complex survey glm regressions.

> data\_analysis\_svy\_adj <- svydesign(id= ~SDMVPSU, strata = ~SDMVSTRA,

+ weights = ~wtmec4yr\_adj\_norm,

+ data = data\_analysis, nest = TRUE)

Error in svydesign(id = ~SDMVPSU, strata = ~SDMVSTRA, weights = ~wtmec4yr\_adj\_norm, :

could not find function "svydesign"

> data\_analysis\_svy\_unadj <- svydesign(id= ~SDMVPSU, strata = ~SDMVSTRA,

+ weights = ~wtmec4yr\_unadj\_norm,

+ data = data\_analysis, nest = TRUE)

Error in svydesign(id = ~SDMVPSU, strata = ~SDMVSTRA, weights = ~wtmec4yr\_unadj\_norm, :

could not find function "svydesign"

>

>

> ## Create empty vector for incrementing included variables in each of the

> ## three weighting procedures

> inc\_vars\_adj <- inc\_vars\_unadj <- inc\_vars\_unwgt <- c()

>

> ## Create empty data frames for storing results for each of the three weighting

> ## procedures. Each of these data frames contains 4 columns:

> ## - Variable: The variable selected in order of inclusion via the forward selection procedure

> ## - Cross-Validated AUC: k-fold cross-valided AUC associated with the forward selection procedure after each variable is selected into the model

> ## - AIC: Complex survey AIC associated with the forward selection procedure after each variable is selected into the model

> ## - AUC: In-sample AUC associated with the forward selection procedure after each variable is selected into the model

> auc\_mat\_adj <- auc\_mat\_unadj <- auc\_mat\_unwgt <-

+ data.frame("Variable" = rep(NA\_character\_,length(ind\_vars)),

+ "Cross-Validated AUC" = rep(NA\_real\_,length(ind\_vars)),

+ "AIC" = rep(NA\_real\_,length(ind\_vars)),

+ "AUC" = rep(NA\_real\_,length(ind\_vars)),

+ stringsAsFactors = FALSE)

>

>

> ## set the seed so cross-validation results are reproducible

> set.seed(1244)

> ## get the training and testing datasets for 10-fold cross validation

> n\_folds <- 10

> ## split the data to have an (approximately) equal number of alive/died in each training/test dataset

> inx\_id\_alive <- which(data\_analysis$yr5\_mort==0)

> inx\_id\_died <- which(data\_analysis$yr5\_mort==1)

> nid\_alive <- length(inx\_id\_alive)

> nid\_died <- length(inx\_id\_died)

>

> inx\_ls\_alive <- split(sample(inx\_id\_alive, size=nid\_alive, replace=FALSE),

+ rep(1:n\_folds,each=ceiling(nid\_alive/n\_folds))[1:nid\_alive])

> inx\_ls\_died <- split(sample(inx\_id\_died, size=nid\_died, replace=FALSE),

+ rep(1:n\_folds,each=ceiling(nid\_died/n\_folds))[1:nid\_died])

> inx\_ls <- lapply(1:n\_folds, function(x) c(inx\_ls\_alive[[x]], inx\_ls\_died[[x]]))

> rm(list=c("inx\_id\_alive","inx\_id\_died","nid\_alive","nid\_died","inx\_ls\_alive","inx\_ls\_died"))

>

>

> ## function which will get cut-points to use in AUC calculation

> ## returns all unique predicted values of the model for a given

> get\_ctpts <- function(model, type='link',...) c(Inf,sort(unique(predict(model,type=type,...)),decreasing=TRUE))

>

> ## Vector to multiply vector of k-fold cross validated weights by to

> ## get an estimate of cross-valided AUC.

> ## This accounts for potentially unequal sizes in the test datasets

> k\_id <- vapply(inx\_ls, length, numeric(1))/nrow(data\_analysis)

>

> for(i in 1:length(ind\_vars)){

+ ## get a vector containing all variables not already included in the forward prediction

+ ## separately for each of the three weighting procedures.

+ exc\_vars\_adj <- setdiff(ind\_vars, inc\_vars\_adj)

+ exc\_vars\_unadj <- setdiff(ind\_vars, inc\_vars\_unadj)

+ exc\_vars\_unwgt <- setdiff(ind\_vars, inc\_vars\_unwgt)

+

+ ## create empty matrices/vectors to store AUC/AIC results

+ ## auc\_ijk\_... Store the cross-valided AUC results

+ ## aic\_ij\_... Store the complex survey AIC for each of the variables not already included in the model

+ ## auc\_ij\_... Store the full sample AUC for each of the variables not already included in the model.

+ auc\_ijk\_adj <- auc\_ijk\_unadj <- auc\_ijk\_unwgt <- matrix(NA, nrow=length(exc\_vars\_adj), ncol=n\_folds)

+ aic\_ij\_adj <- aic\_ij\_unadj <- aic\_ij\_unwgt <-

+ auc\_ij\_full\_adj <- auc\_ij\_full\_unadj <- auc\_ij\_full\_unwgt <- rep(NA,length(exc\_vars\_adj))

+

+

+ ## loop over all variables not currently in the model to calcualte their

+ ## improvment to AUC/AIC.

+ for(j in 1:length(exc\_vars\_adj)){

+ ## current variable to consider

+ var\_adj <- exc\_vars\_adj[j]

+ var\_unadj <- exc\_vars\_unadj[j]

+ var\_unwgt <- exc\_vars\_unwgt[j]

+

+ ## create formulas for current regression using

+ ## all variables previously identified aand the current variable under consideration

+ form\_adj <- paste(c(inc\_vars\_adj, var\_adj), collapse=" +")

+ form\_unadj <- paste(c(inc\_vars\_unadj, var\_unadj), collapse=" +")

+ form\_unwgt <- paste(c(inc\_vars\_unwgt, var\_unwgt), collapse=" +")

+

+

+

+ ## fit the model to the full data for calculating AIC

+ fit\_adj <- svyglm(as.formula(paste("yr5\_mort ~", form\_adj)),

+ design=data\_analysis\_svy\_adj, family=quasibinomial())

+ fit\_unadj <- svyglm(as.formula(paste("yr5\_mort ~", form\_unadj)),

+ design=data\_analysis\_svy\_unadj, family=quasibinomial())

+ fit\_unwgt <- glm(as.formula(paste("yr5\_mort ~", form\_unwgt)), data=data\_analysis, family=binomial())

+

+

+ ## get AIC

+ ## once we have exceeded the design degrees of freedom, we can get

+ ## point estimates for our coefficients, but can no longer calculate AIC

+ if(fit\_adj$df.residual > 0){

+ aic\_ij\_adj[j] <- AIC(fit\_adj)[2]

+ aic\_ij\_unadj[j] <- AIC(fit\_unadj)[2]

+ }

+

+ aic\_ij\_unwgt[j] <- AIC(fit\_unwgt)

+

+

+ ## Calculate weighted and unweighted in-sample AUC using the appropriate weights

+ auc\_ij\_full\_adj[j] <- calc\_weighted\_AUC(response=predict(fit\_adj,type='link'),

+ cutpts=get\_ctpts(fit\_adj,type='link'),

+ labels=data\_analysis$yr5\_mort,

+ weights=data\_analysis$wtmec4yr\_adj\_norm)

+ auc\_ij\_full\_unadj[j] <- calc\_weighted\_AUC(response=predict(fit\_unadj,type='link'),

+ cutpts=get\_ctpts(fit\_unadj,type='link'),

+ labels=data\_analysis$yr5\_mort,

+ weights=data\_analysis$wtmec4yr\_unadj\_norm)

+ auc\_ij\_full\_unwgt[j] <- calc\_weighted\_AUC(response=predict(fit\_unwgt,type='link'),

+ cutpts=get\_ctpts(fit\_unwgt,type='link'),

+ labels=data\_analysis$yr5\_mort,

+ weights=rep(1, nrow(data\_analysis)))

+

+ ## get cross-validated AUC

+ for(k in 1:n\_folds){

+ ## subset test and training data sets

+ SEQN\_train <- data\_analysis$SEQN[-inx\_ls[[k]]]

+ SEQN\_test <- data\_analysis$SEQN[inx\_ls[[k]]]

+ data\_test <- subset(data\_analysis, SEQN %in% SEQN\_test)

+ data\_train <- subset(data\_analysis, SEQN %in% SEQN\_train)

+

+

+ ## Fit the appropriate models by subsetting the data.

+ ## By subsetting the existing svydesign objects instead of creating new svydesign objects,

+ ## we retain information on the number of PSU/strata in the original study.

+ fit\_adj\_cv <- svyglm(as.formula(paste("yr5\_mort ~", form\_adj)),

+ design=subset(data\_analysis\_svy\_adj, SEQN %in% SEQN\_train),

+ family=quasibinomial())

+ fit\_unadj\_cv <- svyglm(as.formula(paste("yr5\_mort ~", form\_unadj)),

+ design=subset(data\_analysis\_svy\_unadj, SEQN %in% SEQN\_train),

+ family=quasibinomial())

+ fit\_unwgt\_cv <- glm(as.formula(paste("yr5\_mort ~", form\_unwgt)), data=data\_train,

+ family=quasibinomial())

+

+

+ ## Calculate weighted and unweighted AUC using the appropriate weights

+ auc\_ijk\_adj[j,k] <- calc\_weighted\_AUC(response=predict(fit\_adj\_cv, newdata=data\_test, type='link'),

+ cutpts=get\_ctpts(fit\_adj\_cv,type='link',newdata=data\_test),

+ labels=data\_test$yr5\_mort,

+ weights=data\_test$wtmec4yr\_adj\_norm)

+ auc\_ijk\_unadj[j,k] <- calc\_weighted\_AUC(response=predict(fit\_unadj\_cv, newdata=data\_test, type='link'),

+ cutpts=get\_ctpts(fit\_unadj\_cv,type='link',newdata=data\_test),

+ labels=data\_test$yr5\_mort,

+ weights=data\_test$wtmec4yr\_unadj\_norm)

+ auc\_ijk\_unwgt[j,k] <- calc\_weighted\_AUC(response=predict(fit\_unwgt\_cv, newdata=data\_test, type='link'),

+ cutpts=get\_ctpts(fit\_unwgt\_cv,type='link',newdata=data\_test),

+ labels=data\_test$yr5\_mort,

+ weights=rep(1,nrow(data\_test)))

+

+

+ rm(list=c("data\_train","data\_test","SEQN\_train","SEQN\_test",

+ paste0("fit\_",c("adj","unadj","unwgt") ,"\_cv")))

+ }

+

+

+

+ print(j)

+ rm(list=c(paste0("fit\_",c("adj","unadj","unwgt")),

+ paste0("form\_",c("adj","unadj","unwgt")),

+ paste0("var\_",c("adj","unadj","unwgt")),

+ "k"))

+ }

+

+ ## average across the k-folds using a weighted average of the number of individuals in each test set

+ auc\_ij\_adj <- auc\_ijk\_adj %\*% k\_id

+ auc\_ij\_unadj <- auc\_ijk\_unadj %\*% k\_id

+ auc\_ij\_unwgt <- auc\_ijk\_unwgt %\*% k\_id

+

+ ## combine results for this iteratin

+ auc\_j\_adj <- data.frame(exc\_vars\_adj, auc\_ij\_adj, aic\_ij\_adj, auc\_ij\_full\_adj, stringsAsFactors = FALSE)

+ auc\_j\_unadj <- data.frame(exc\_vars\_unadj, auc\_ij\_unadj, aic\_ij\_unadj, auc\_ij\_full\_unadj, stringsAsFactors = FALSE)

+ auc\_j\_unwgt <- data.frame(exc\_vars\_unwgt, auc\_ij\_unwgt, aic\_ij\_unwgt, auc\_ij\_full\_unwgt, stringsAsFactors = FALSE)

+

+ ## identify which variable is associated with the best improvement (or least bad decrease) in CV-AUC

+ auc\_mat\_adj[i,] <- auc\_j\_adj[which.max(auc\_j\_adj[,2]),]

+ auc\_mat\_unadj[i,] <- auc\_j\_unadj[which.max(auc\_j\_unadj[,2]),]

+ auc\_mat\_unwgt[i,] <- auc\_j\_unwgt[which.max(auc\_j\_unwgt[,2]),]

+

+ ## Create matrices which communicate the predictive power of all univariate regressions.

+ ## These are used to create Table 4 in the manuscript.

+ if(i == 1){

+ auc\_mat\_1\_adj <- auc\_j\_adj[rev(order(auc\_j\_adj[,2])),]

+ auc\_mat\_1\_unadj <- auc\_j\_unadj[rev(order(auc\_j\_unadj[,2])),]

+ auc\_mat\_1\_unwgt <- auc\_j\_unwgt[rev(order(auc\_j\_unwgt[,2])),]

+ }

+

+ ## Add in our "best" variable to our list of variabels included in the analysis

+ inc\_vars\_adj <- c(inc\_vars\_adj, auc\_mat\_adj[i,1])

+ inc\_vars\_unadj <- c(inc\_vars\_unadj, auc\_mat\_unadj[i,1])

+ inc\_vars\_unwgt <- c(inc\_vars\_unwgt, auc\_mat\_unwgt[i,1])

+

+

+ ## at the end of each iteration, print out the current

+ ## adjusted survey weight results

+ print(paste(i, " Independent variable finished"))

+ print(auc\_mat\_adj[1:i,]);

+ rm(list=c(paste0("auc\_j\_",c("adj","unadj","unwgt")),

+ paste0("auc\_ij\_",c("adj","unadj","unwgt")),

+ paste0("auc\_ijk\_",c("adj","unadj","unwgt")),

+ paste0("auc\_ij\_full\_",c("adj","unadj","unwgt")),

+ paste0("aic\_ij\_",c("adj","unadj","unwgt")),

+ "j"))

+ }

Error in svyglm(as.formula(paste("yr5\_mort ~", form\_adj)), design = data\_analysis\_svy\_adj, :

could not find function "svyglm"

> rm(list=c("inx\_ls",

+ paste0("inc\_vars\_",c("adj","unadj","unwgt")),

+ paste0("exc\_vars\_",c("adj","unadj","unwgt")),

+ "i","n\_folds","k\_id","ind\_vars"))

>

>

>

> ## Final model fits using the adjusted survey weights.

> ## Fit models using the first 7 predictors (where all models agree) as well as

> ## the number of predictors identified by AIC and AUC criteria.

> inx\_svy\_adj <- 7

> inx\_aic\_svy\_adj <- min(which(diff(auc\_mat\_adj[,3]) > 0))

Warning message:

In min(which(diff(auc\_mat\_adj[, 3]) > 0)) :

no non-missing arguments to min; returning Inf

> inx\_auc\_svy\_adj <- min(which(diff(auc\_mat\_adj[,2]) < 0))

Warning message:

In min(which(diff(auc\_mat\_adj[, 2]) < 0)) :

no non-missing arguments to min; returning Inf

>

> fit\_final <- svyglm(as.formula(paste("yr5\_mort ~", paste0(auc\_mat\_adj[1:inx\_svy\_adj,1] ,collapse="+") )),

+ design=data\_analysis\_svy\_adj,

+ family=quasibinomial())

Error in svyglm(as.formula(paste("yr5\_mort ~", paste0(auc\_mat\_adj[1:inx\_svy\_adj, :

could not find function "svyglm"

> fit\_final\_aic <- svyglm(as.formula(paste("yr5\_mort ~", paste0(auc\_mat\_adj[1:inx\_aic\_svy\_adj,1] ,collapse="+") )),

+ design=data\_analysis\_svy\_adj,

+ family=quasibinomial())

Error in svyglm(as.formula(paste("yr5\_mort ~", paste0(auc\_mat\_adj[1:inx\_aic\_svy\_adj, :

could not find function "svyglm"

> fit\_final\_auc <- svyglm(as.formula(paste("yr5\_mort ~", paste0(auc\_mat\_adj[1:inx\_auc\_svy\_adj,1] ,collapse="+") )),

+ design=data\_analysis\_svy\_adj,

+ family=quasibinomial())

Error in svyglm(as.formula(paste("yr5\_mort ~", paste0(auc\_mat\_adj[1:inx\_auc\_svy\_adj, :

could not find function "svyglm"

>

> ## look at the coefficients for the first 7 coefficients

> ## selected using the adjusted survey weight model

> summary(fit\_final)

Call:

svyglm(formula = as.formula(paste("yr5\_mort ~", paste0(auc\_mat\_adj[1:inx\_svy\_adj,

1], collapse = "+"))), design = data\_analysis\_svy\_adj, family = quasibinomial())

Deviance Residuals:

Min 1Q Median 3Q Max

-1.5680 -0.3936 -0.2263 -0.1163 5.5006

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -6.03571 12.23431 -0.493 0.628

TAC -0.71854 1.92476 -0.373 0.713

Age 0.06495 0.14266 0.455 0.654

GenderFemale -0.81124 2.44069 -0.332 0.743

MobilityProblemAny Difficulty 0.81220 2.53162 0.321 0.752

DrinkStatusNon-Drinker 0.71991 2.60548 0.276 0.785

DrinkStatusHeavy Drinker 0.74208 4.42667 0.168 0.869

DrinkStatusMissing alcohol 0.92570 5.17516 0.179 0.860

SmokeCigsFormer 0.50059 2.77084 0.181 0.859

SmokeCigsCurrent 0.87292 3.49667 0.250 0.806

BMI\_catUnderweight 0.65432 8.18681 0.080 0.937

BMI\_catOverweight -0.69810 2.92541 -0.239 0.814

BMI\_catObese -0.55701 2.94961 -0.189 0.852

(Dispersion parameter for quasibinomial family taken to be 253.6167)

Null deviance: 1771.6 on 3197 degrees of freedom

Residual deviance: 1366.4 on 18 degrees of freedom

AIC: NA

Number of Fisher Scoring iterations: 6

>

>

> ####################################################

> # Create Tables: #

> # 1) single best predictor ordered by AUC #

> # 2) forward selection procedure #

> # 3) final mode coefficient estimates and 95% CI #

> ####################################################

> if(make\_tables){

+ source(file.path(code\_path, "create\_table\_auc.R"))

+

+ ## sourcing "create\_table\_final\_regression.R" will create one warning which can be safely ignored

+ source(file.path(code\_path, "create\_table\_final\_regression.R"))

+ }

--------------------------------------------------------------------------------------------

Single predictor Cross-validated AUC

--------------------------------------------------------------------------------------------

[,1] [,2] [,3]

[1,] "1" "TAC" "0.781"

[2,] "2" "MVPA" "0.755"

[3,] "3" "Sedentary, Sleep, or Non-wear" "0.744"

[4,] "4" "ASTP$\_{sl/nw}$" "0.743"

[5,] "5" "Age" "0.741"

[6,] "6" "TLAC" "0.738"

[7,] "7" "Mobility problem" "0.676"

[8,] "8" "SATP$\_{sl/nw}$" "0.675"

[9,] "9" "Surrogate for $s\_{i6}$" "0.663"

[10,] "10" "Alcohol consumption" "0.608"

[11,] "11" "Education" "0.599"

[12,] "12" "Surrogate for $s\_{i1}$" "0.582"

[13,] "13" "Cancer" "0.581"

[14,] "14" "Cigarette smoking" "0.576"

[15,] "15" "Congestive heart failure" "0.567"

[16,] "16" "Gender" "0.563"

[17,] "17" "Body mass index" "0.561"

[18,] "18" "Diabetes" "0.552"

[19,] "19" "Coronary heart disease" "0.547"

[20,] "20" "Stroke" "0.542"

[21,] "21" "Race" "0.492"

[22,] "22" "Wear time" "0.468"

[,4] [,5] [,6]

[1,] "TAC" "0.782" "TAC"

[2,] "MVPA" "0.755" "Age"

[3,] "Sedentary, Sleep, or Non-wear" "0.746" "MVPA"

[4,] "ASTP$\_{sl/nw}$" "0.743" "ASTP$\_{sl/nw}$"

[5,] "TLAC" "0.740" "Sedentary, Sleep, or Non-wear"

[6,] "Age" "0.739" "TLAC"

[7,] "SATP$\_{sl/nw}$" "0.676" "SATP$\_{sl/nw}$"

[8,] "Mobility problem" "0.676" "Mobility problem"

[9,] "Surrogate for $s\_{i6}$" "0.662" "Surrogate for $s\_{i6}$"

[10,] "Alcohol consumption" "0.605" "Gender"

[11,] "Education" "0.600" "Alcohol consumption"

[12,] "Surrogate for $s\_{i1}$" "0.584" "Cigarette smoking"

[13,] "Cancer" "0.581" "Surrogate for $s\_{i1}$"

[14,] "Cigarette smoking" "0.578" "Cancer"

[15,] "Congestive heart failure" "0.568" "Education"

[16,] "Gender" "0.566" "Congestive heart failure"

[17,] "Body mass index" "0.562" "Race"

[18,] "Diabetes" "0.552" "Stroke"

[19,] "Coronary heart disease" "0.548" "Diabetes"

[20,] "Stroke" "0.541" "Coronary heart disease"

[21,] "Race" "0.492" "Body mass index"

[22,] "Wear time" "0.463" "Wear time"

[,7]

[1,] "0.751"

[2,] "0.732"

[3,] "0.729"

[4,] "0.726"

[5,] "0.723"

[6,] "0.715"

[7,] "0.654"

[8,] "0.653"

[9,] "0.652"

[10,] "0.582"

[11,] "0.575"

[12,] "0.573"

[13,] "0.571"

[14,] "0.567"

[15,] "0.567"

[16,] "0.558"

[17,] "0.550"

[18,] "0.541"

[19,] "0.541"

[20,] "0.540"

[21,] "0.539"

[22,] "0.495"

---------------------------------------------------------------------------------------------

Forward Selection AUC

--------------------------------------------------------------------------------------------

[,1] [,2] [,3] [,4] [,5] [,6]

[1,] NA " NA" NA " NA" NA " NA"

[2,] NA " NA" NA " NA" NA " NA"

[3,] NA " NA" NA " NA" NA " NA"

[4,] NA " NA" NA " NA" NA " NA"

[5,] NA " NA" NA " NA" NA " NA"

[6,] NA " NA" NA " NA" NA " NA"

[7,] NA " NA" NA " NA" NA " NA"

[8,] NA " NA" NA " NA" NA " NA"

[9,] NA " NA" NA " NA" NA " NA"

[10,] NA " NA" NA " NA" NA " NA"

[11,] NA " NA" NA " NA" NA " NA"

[12,] NA " NA" NA " NA" NA " NA"

[13,] NA " NA" NA " NA" NA " NA"

[14,] NA " NA" NA " NA" NA " NA"

[15,] NA " NA" NA " NA" NA " NA"

[16,] NA " NA" NA " NA" NA " NA"

[17,] NA " NA" NA " NA" NA " NA"

[18,] NA " NA" NA " NA" NA " NA"

[19,] NA " NA" NA " NA" NA " NA"

[20,] NA " NA" NA " NA" NA " NA"

[21,] NA " NA" NA " NA" NA " NA"

[22,] NA " NA" NA " NA" NA " NA"

Error in kable(auc\_1, format = "latex", booktabs = TRUE, escape = FALSE) :

could not find function "kable"

In addition: There were 18 warnings (use warnings() to see them)

>

>

>

>

>