```
/*

PHD PROJECT: The role of depressive symptoms and cardiometabolic risk factors in the prediction of dementia: a cross-country comparison in England, the United States and China

STUDY I: Independent and combined effects of depressive symptoms and cardiometabolic risk factors on dementia incidence

DATASET: HRS
baseline: wave 8 (2004) follow-up waves 9-14 (2008-2018)



TIMELINE

DEPRESSIVE SYMPTOMS AND CARDIOMETABOLIC RISK FACTORS: WV8 (BASELINE)
DEMENTIA INCIDENCE: W9 - WV14 (6 TIME POINTS)
COVARIATES ADJUSTMENT FOR HR MODELS: WV8

*/


* KEEP NECESSARY VARIABLES


keep HHID PN RAHHIDPN HHIDPN ///
H_sex H_age H_eduaction_yrs H_eduaction H_maritalstatus_4cat H_wealthquintiles ///
H_ethnicity H_hispanic_ethnicity ///
H_smoking_2cat H_smoking_3cat H_physicalactivity H_alcohol_freq H_alcohol_status ///
H_cvd_comorbidity Hwv8_cognition Hwv8_memory Hwv8_loneliness_quintiles Hwv12_memory ///
Hwv8_cesd_sumscore Hwv8_depressive_symptoms ///
Hwv9_cesd_sumscore Hwv9_depressive_symptoms ///
Hwv10_cesd_sumscore Hwv10_depressive_symptoms ///
Hwv11_cesd_sumscore Hwv11_depressive_symptoms ///
Hwv12_cesd_sumscore Hwv12_depressive_symptoms ///
Hwv13_cesd_sumscore Hwv13_depressive_symptoms ///
Hwv14_cesd_sumscore Hwv14_depressive_symptoms ///
Hwv8_crp_level Hwv8_crp Hwv8_hdl_level Hwv8_male_hdl Hwv8_female_hdl ///
Hwv8_meds_hdl Hwv8_hdl_sum Hwv8_hdl ///
Hwv8_waist Hwv8_malewaist_ao Hwv8_femalewaist_ao Hwv8_obesity_waist_sum Hwv8_obesity_waist ///
Hwv8_bmi_score Hwv8_obesity_bmi Hwv8_waist_bmi_sum Hwv8_obesity ///
Hwv8_systolic_mean Hwv8_diastolic_mean Hwv8_systolic_bp Hwv8_diastolic_bp ///
Hwv8_meds_bp Hwv8_bp_before Hwv8_bp_report Hwv8_bpevr ///
Hwv8_bp_reportevr_sum Hwv8_bp_reportevr Hwv8_bp_sum Hwv8_bp ///
Hwv8_diabetes_before Hwv8_diabetes_report Hwv8_diabetesevr ///
Hwv8_diabetes_reportevr_sum Hwv8_diabetes_reportevr ///
Hwv8_meds_diabetes Hwv8_insulin_diabetes ///
Hwv8_diabetes_anymeds_sum Hwv8_diabetes_anymeds ///
Hwv8_HbA1c_level Hwv8_HbA1c Hwv8_diabetes_HbA1c_sum Hwv8_glycemia ///
Hwv8_ao_depress_sum Hwv8_Nao_Ndepress Hwv8_Nao_Ydepress ///
Hwv8_Yao_Ndepress Hwv8_Yao_Ydepress Hwv8_ao_depress_group ///
Hwv8_waist_depress_sum Hwv8_Nwaist_Ndepress Hwv8_Nwaist_Ydepress ///
Hwv8_Ywaist_Ndepress Hwv8_Ywaist_Ydepress Hwv8_waist_depress_group ///
Hwv8_glycemia_depress_sum Hwv8_Nglycemia_Ndepress ///
Hwv8_Nglycemia_Ydepress Hwv8_Yglycemia_Ndepress ///
Hwv8_Yglycemia_Ydepress Hwv8_glycemia_depress_group ///
Hwv8_diabet_depress_sum Hwv8_Ndiabet_Ndepress ///
Hwv8_Ndiabet_Ydepress Hwv8_Ydiabet_Ndepress ///
Hwv8_Ydiabet_Ydepress Hwv8_diabet_depress_group ///
Hwv8_hba1c_depress_sum Hwv8_Nhba1c_Ndepress ///
Hwv8_Nhba1c_Ydepress Hwv8_Yhba1c_Ndepress Hwv8_Yhba1c_Ydepress ///
Hwv8_hba1c_depress_group Hwv8_hdl_depress_sum Hwv8_Nhdl_Ndepress ///
Hwv8_Nhdl_Ydepress Hwv8_Yhdl_Ndepress Hwv8_Yhdl_Ydepress ///
Hwv8_hdl_depress_group Hwv8_bp_depress_sum Hwv8_Nbp_Ndepress ///
Hwv8_Nbp_Ydepress Hwv8_Ybp_Ndepress Hwv8_Ybp_Ydepress ///
Hwv8_bp_depress_group Hwv8_sbp_depress_sum Hwv8_Nsbp_Ndepress ///
Hwv8_Nsbp_Ydepress Hwv8_Ysbp_Ndepress Hwv8_Ysbp_Ydepress ///
Hwv8_sbp_depress_group Hwv8_dbp_depress_sum Hwv8_Ndbp_Ndepress ///
```

```
Hwv8_Ndbp_Ydepress Hwv8_Ydbp_Ndepress Hwv8_Ydbp_Ydepress ///
Hwv8_dbp_depress_group Hwv8_crp_depress_sum Hwv8_Ncrp_Ndepress ///
Hwv8_Ncrp_Ydepress Hwv8_Ycrp_Ndepress Hwv8_Ycrp_Ydepress ///
Hwv8_crp_depress_group Hwv8_cardio3_sum Hwv8_cardio3 ///
Hwv8_ca3_depress_sum Hwv8_Nca3_Ndepress Hwv8_Nca3_Ydepress ///
Hwv8_Yca3_Ndepress Hwv8_Yca3_Ydepress Hwv8_ca3_depress_group ///
Hwv8_cardio4_sum Hwv8_cardio4 Hwv8_ca4_depress_sum ///
Hwv8_Nca4_Ndepress Hwv8_Nca4_Ydepress Hwv8_Yca4_Ndepress ///
Hwv8_Yca4_Ydepress Hwv8_ca4_depress_group ///
Hwv8_cardio_number_sum Hwv8_cardio_number ///
Hwv8_cardio2_sum Hwv8_cardio2 Hwv8_ca2_depress_sum ///
Hwv8_Nca2_Ndepress Hwv8_Nca2_Ydepress ///
Hwv8_Yca2_Ndepress Hwv8_Yca2_Ydepress Hwv8_ca2_depress_group ///
Hwv8_memory_report Hwv9_memory_report Hwv10_anydementia_report ///
Hwv11_anydementia_report Hwv12_anydementia_report ///
Hwv13_anydementia_report Hwv14_anydementia_report ///
Hwv8_interview_date Hwv9_interview_date Hwv10_interview_date ///
Hwv11_interview_date Hwv12_interview_date Hwv13_interview_date Hwv14_interview_date ///
Hwv9to14_dementia_sum Hwv9to14_dementia_event ///
Hwv9to14_newdementia_or_lastinte Hwv9to14_dementia_free_date H_time_dementia_months ///
H_time_dementia_midpoint H_time_dementia_midpoint_final H_time_of_event_dementia ///
Hwv12to14_newdementia_or_lastint Hwv12to14_time_dementia_months ///
Hwv12to14_dementia_free_date Hwv12to14_time_dementia_midpoint ///
Hwv12to14_time_dementia_midpoin0 Hwv12to14_time_of_event_dementia



/* ---- MERGE DATA ----

Process to merge

Open master dataset and run merge two datasets

After merging all data both from master and using will be added
Need to keep if _merge==3
1 means cases from master data
2 means cases from using data
3 means cases from both master and using data

Drop _merge var.
*/


help merge

* Menu > Data > Combine datasets > Merge two datasets
* Choose One to many (key variable)


merge 1:m RAHHIDPN using
"S:\Research\pkstudies\Study1_biopsych_risk\HRS\hrs_to_merge_physicalact.dta"

keep if _merge==3

drop _merge




/*

EXPOSURE VARIABLES

Binary variables of depressive symptoms and cardiometabolic markers measured at wave 8


Depression: Hwv8_depressive_symptoms
```

```
134
135   CRP: Hwv8_crp
136
137   HDL cholesterol: Hwv8_hdl
138
139   Obesity by waist cir: Hwv8_obesity_waist
140
141   systolic Blood pressure: Hwv8_systolic_bp
142
143   diastolic Blood pressure: Hwv8_diastolic_bp
144
145   Diabetes: Hwv8_diabetes_reportevr
146
147   HbA1c: Hwv8_HbA1c
148
149
150   CA number (categ 0,1,2,3,4+): Hwv8_cardio_number
151
152   CA mutlimorbidity >= 2 CA conditions
153
154   Grouping of Dep-CA:
155   Hwv8_waist_depress_group Hwv8_diabet_depress_group Hwv8_hba1c_depress_group
      Hwv8_hdl_depress_group Hwv8_sbp_depress_group Hwv8_dbp_depress_group Hwv8_crp_depress_group
      Hwv8_ca3_depress_group Hwv8_ca4_depress_group Hwv8_ca2_depress_group
156
157
158   OUTCOME VARIABLES
159
160   Dementia event: Hwv9to14_dementia_event
161
162   Time-to-event: H_time_of_event_dementia
163
164
165   */
166
167
168
169
170
171
172   *** Descriptive stats of var of interest
173
174
175   tabulate Hwv8_depressive_symptoms
176   summarize Hwv8_depressive_symptoms
177
178   misstable summarize Hwv8_depressive_symptoms
179   misstable patterns Hwv8_depressive_symptoms
180
181
182   tabulate Hwv8_crp
183   summarize Hwv8_crp
184
185   misstable summarize Hwv8_crp
186   misstable patterns Hwv8_crp
187
188   tabulate Hwv8_hdl
189   summarize Hwv8_hdl
190
191   misstable summarize Hwv8_hdl
192   misstable patterns Hwv8_hdl
193
194   tabulate Hwv8_obesity_waist
195   summarize Hwv8_obesity_waist
196
197   misstable summarize Hwv8_obesity_waist
198   misstable patterns Hwv8_obesity_waist
199
```

```
tabulate Hwv8_systolic_bp
summarize Hwv8_systolic_bp

misstable summarize Hwv8_systolic_bp
misstable patterns Hwv8_systolic_bp


tabulate Hwv8_diastolic_bp
summarize Hwv8_diastolic_bp

misstable summarize Hwv8_diastolic_bp
misstable patterns Hwv8_diastolic_bp


tabulate Hwv8_diabetes_reportevr
summarize Hwv8_diabetes_reportevr

misstable summarize Hwv8_diabetes_reportevr
misstable patterns Hwv8_diabetes_reportevr


tabulate Hwv8_HbA1c
summarize Hwv8_HbA1c

misstable summarize Hwv8_HbA1c
misstable patterns Hwv8_HbA1c


tabulate Hwv8_memory_report
summarize Hwv8_memory_report

misstable summarize Hwv8_memory_report
misstable patterns Hwv8_memory_report









*** CLEANING DATA

* 1. drop dementia cases and missing data at baseline

* drop dementia wave 8 missing data
drop if Hwv8_memory_report==1
* (226 observations deleted)

drop if Hwv8_memory_report== .
* (0 observations deleted)



* 2. drop missing values and invalid data of depresssive sym and cardiometabolic markers


drop if Hwv8_depressive_symptoms== .
* (44 observations deleted)

drop if Hwv8_crp== .
* (499 observations deleted)

drop if Hwv8_crp_level > 100 & Hwv8_crp_level < 300
* (1 observations deleted)


drop if Hwv8_hdl== .
* (225 observations deleted)

```

```
drop if Hwv8_obesity_waist== .
* (188 observations deleted)

drop if Hwv8_systolic_bp== .
* (102 observations deleted)

* drop SBP > 300

drop if Hwv8_systolic_mean > 300 & Hwv8_systolic_mean < 1000
* (81 observations deleted)

drop if Hwv8_diastolic_bp== .
* (0 observations deleted)

* drop DBP > 300

drop if Hwv8_diastolic_mean > 300 & Hwv8_diastolic_mean < 1000
* (17 observations deleted)


drop if Hwv8_diabetes_reportevr== .
* (4 observations deleted)

drop if Hwv8_HbA1c== .
* (74 observations deleted)



* 3. drop obs with no records on dementia at any wave from 9-14 follow-ups


search mdesc
search rmiss2
search mvpatterns

* see number of missing values vs non-missing in each variable
mdesc Hwv9_memory_report Hwv10_anydementia_report Hwv11_anydementia_report ///
Hwv12_anydementia_report Hwv13_anydementia_report Hwv14_anydementia_report



/* number of missing values per observation
* the code below creates a variable called nmisfollowup that gives the number of missing values
for each observation in the variables of interest */
egen nmisfollowup_dementia_wv9to14=rmiss2(Hwv9_memory_report ///
Hwv10_anydementia_report Hwv11_anydementia_report ///
Hwv12_anydementia_report Hwv13_anydementia_report Hwv14_anydementia_report)

tab nmisfollowup_dementia_wv9to14

* drop observations "nmisfollowup_dementia_wv9to14" > 5 (those with 6 missing data = no records
at any wave)
drop if nmisfollowup_dementia_wv9to14>5
*(253 observations deleted)


* ANALYTIC SAMPLE -> 5021









/*
---- DESCRIPTIVE STATISTICS ----

General characteristics of participnats stratified for study inclusion
```

```
335
336 General characteristics of participants stratified for dementia occurence
337
338 1. CHI-SQUARE (chi2) for categorical var (crosstabulation)
339     Frequency tables -> two-way tables
340         using the command tabulate, chi2
341         reporting observations, column percentage (N, %) and p-value of Pearson's r
342
343
344 2. one-way ANOVA for continuous var
345     check box plot
346     using the command oneway
347     reporting mean, sd (summary tables) and p-value of F
348 */
349
350
351
352
353 * General characteristics of HRS participants at baseline
354
355
356 * Socio-demographics
357 sum H_age
358 ta H_sex
359 ta H_eduaction
360 ta H_maritalstatus_4cat
361 ta H_wealthquintiles
362 * Cardiometabolic risk factors
363 ta Hwv8_crp
364 ta Hwv8_hdl
365 ta Hwv8_obesity_waist
366 ta Hwv8_systolic_bp
367 ta Hwv8_diastolic_bp
368 ta Hwv8_diabetes_reportevr
369 ta Hwv8_HbA1c
370 ta Hwv8_cardio2
371 * Lifestyle and health indicators
372 ta H_smoking_3cat
373 ta H_physicalactivity
374 ta H_alcohol_status
375 ta H_cvd_comorbidity
376 * Depressive symptoms
377 ta Hwv8_depressive_symptoms
378 * Memory score
379 sum Hwv8_memory
380
381
382
383
384 * General characteristics of HRS participants stratified for dementia occurence
385
386 * Socio-demographics
387 ttest H_age, by(Hwv9to14_dementia_event)
388 ta H_sex Hwv9to14_dementia_event, chi2 column row
389 ta H_eduaction Hwv9to14_dementia_event, chi2 column row
390 ta H_maritalstatus_4cat Hwv9to14_dementia_event, chi2 column row
391 ta H_wealthquintiles Hwv9to14_dementia_event, chi2 column row
392 * Cardiometabolic risk factors
393 ta Hwv8_crp Hwv9to14_dementia_event, chi2 column row
394 ta Hwv8_hdl Hwv9to14_dementia_event, chi2 column row
395 ta Hwv8_obesity_waist Hwv9to14_dementia_event, chi2 column row
396 ta Hwv8_systolic_bp Hwv9to14_dementia_event, chi2 column row
397 ta Hwv8_diastolic_bp Hwv9to14_dementia_event, chi2 column row
398 ta Hwv8_diabetes_reportevr Hwv9to14_dementia_event, chi2 column row
399 ta Hwv8_HbA1c Hwv9to14_dementia_event, chi2 column row
400 ta Hwv8_cardio2 Hwv9to14_dementia_event, chi2 column row
401 * Lifestyle and health indicators
402 ta H_smoking_3cat Hwv9to14_dementia_event, chi2 column row
```

```
ta H_physicalactivity Hwv9to14_dementia_event, chi2 column row
ta H_alcohol_status Hwv9to14_dementia_event, chi2 column row
ta H_cvd_comorbidity Hwv9to14_dementia_event, chi2 column row
* Depressive symptoms
ta Hwv8_depressive_symptoms Hwv9to14_dementia_event, chi2 column row
* Memory score
ttest Hwv8_memory, by(Hwv9to14_dementia_event)
ta H_age_group Hwv9to14_dementia_event, chi2 column row









/*
---- SURVIVAL ANALYSIS IN COMPLETE DATA ----

Tests of proportional-hazards assumption
Kaplan Meier survival curves
Person-time
Cox proportional regression - Hazard ratios - stcox
Postestimation tools for stcox
Test of Goodness of Fit

*** Cox regression in full data, complete data (listwise deletion of missing data) and imputed data
Cox PH regression in complete data
Cox PH regression model in imputed dataset - mi estimate


*/



* check dataset variables of interest only

codebook H_time_of_event_dementia Hwv9to14_dementia_event ///
Hwv8_depressive_symptoms  Hwv8_crp ///
Hwv8_hdl Hwv8_obesity_waist Hwv8_systolic_bp Hwv8_diastolic_bp ///
Hwv8_diabetes_reportevr Hwv8_HbA1c ///
Hwv8_cardio3 Hwv8_cardio4 ///
Hwv8_cardio_number_sum Hwv8_cardio_number ///
Hwv8_waist_depress_group Hwv8_diabet_depress_group ///
Hwv8_hba1c_depress_group Hwv8_hdl_depress_group ///
Hwv8_sbp_depress_group Hwv8_dbp_depress_group ///
Hwv8_crp_depress_group ///
Hwv8_ca3_depress_group Hwv8_ca4_depress_group ///
H_age H_sex H_eduaction H_maritalstatus_4cat H_wealthquintiles ///
H_smoking_3cat H_alcohol_status H_cvd_comorbidity,compact



* Declare Data to be Survival Data
* Time to event: H_time_of_event_dementia (months)
* Censoring: Hwv9to14_dementia_event (1=dementia, 0=censored)
* Command is stset TIMETOEVENT, failure(CENSORVARIABLE)


stset H_time_of_event_dementia, failure (Hwv9to14_dementia_event==1) id(RAHHIDPN)



*describe survival data using commnad stsum

stsum

stsum, by(Hwv8_ca2_depress_group)
```

```
471
472
473   * Kaplan Meier Curve estimation
474
475   sts list
476
477   sts list, by(Hwv8_ca2_depress_group)
478
479
480
481   * Kaplan Meier Curve Plot
482
483   * no frills plot
484
485   sts graph
486
487   * with frills
488
489   sts graph, xtitle("Time in Months") ytitle("Survival Prob") ///
490   title("Kaplan Meier Curve")
491
492
493   * With Greenwood CI limits
494
495   sts graph, gwood legend(off) xtitle("Time in Months") ytitle("Survival Prob") ///
496   title("Kaplan Meier Curve")
497
498
499
500   * Group Kaplan-Meier Curve Estimation
501   * Command is sts graph, by(GROUPVAR) OPTION OPTION OPTION Note: Must have sorted by GROUPVAR first
502
503   sort Hwv8_ca2_depress_group
504
505   sts list, by(Hwv8_ca2_depress_group)
506
507   * graph with frills
508
509   sts graph, by(Hwv8_ca2_depress_group) xlabel(0(20)180) ylabel(0.80(.05)1) xtitle("Time in Months")
       ///
510   ytitle("Survival Prob") title("Kaplan Meier Curve")
511
512
513
514
515   * calculate person-time and incidence rates using command ststime
516
517   stptime,title(Person-years)
518
519   stptime, title(Person-years) per(1000)
520
521
522   /*
523
524   Repeat to find incident case per category
525
526   Hwv8_depressive_symptoms
527   Hwv8_crp
528   Hwv8_hdl
529   Hwv8_obesity_waist
530   Hwv8_systolic_bp
531   Hwv8_diastolic_bp
532   Hwv8_diabetes_reportevr
533   Hwv8_HbA1c
534   Hwv8_cardio_number
535   Hwv8_cardio2
536
537   Hwv8_crp_depress_group
```

```
Hwv8_hdl_depress_group
Hwv8_waist_depress_group
Hwv8_sbp_depress_group
Hwv8_dbp_depress_group
Hwv8_diabet_depress_group
Hwv8_hba1c_depress_group
Hwv8_ca2_depress_group

*/



ta Hwv8_ca2_depress_group

* calculate person-time  by category

stptime, by(Hwv8_ca2_depress_group)

stptime, by(Hwv8_ca2_depress_group) per(1000)



* mean and median of follow-up
sum H_time_of_event_dementia
sum H_time_of_event_dementia, detail




/* Log Rank Test of equality of survival distributions
 (NULL: equality of survival distributions among groups)
 We will consider including the predictor if the test has a p-value of 0.2 – 0.25 or less.
 If the predictor has a p-value greater than 0.25 in a univariate analysis
 it is highly unlikely that it will contribute anything to a model which includes other predictors.
 Command is sts test GROUPVAR
*/


sts test Hwv8_cardio2, logrank

sts test Hwv8_ca2_depress_group, logrank

sts test H_age, logrank

sts test H_sex, logrank

sts test H_eduaction, logrank

sts test H_maritalstatus_4cat, logrank

sts test H_wealthquintiles, logrank

sts test H_smoking_3cat, logrank

sts test H_alcohol_status, logrank

sts test H_cvd_comorbidity, logrank




/* Cox PH regression model

using the command stcox
```

```
605
606   --- Building the model ---
607
608   Model 1: unadjusted - single predictor of CM classes
609   Model 2: model 1 + sociodemographics: age sex education marital status and wealth
610   Model 3: model 2 + lifestyle/health indicators: smoking, alcohol consumption, cvd comorbidity
611
612
613
614   !! I didn't adjust for physical activity because this variable can't be used in CHARLS (missing
      values)
615
616   */
617
618
619   * Unadjusted model - model 1 - single predictor
620
621   stcox Hwv8_ca2_depress_group
622
623   * define design var by using i.(by group)
624
625   stcox i.Hwv8_ca2_depress_group
626
627
628   * Adjusted models - multivariable Cox model
629   * controlling for covariates
630
631   * model 2: model 1 + adjust for socio-demographics: age sex education marital status and wealth
632
633   stcox i.Hwv8_ca2_depress_group H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.
      H_wealthquintiles
634
635
636   * model 3: model 2 + adjust for lifestyle/ health indicators
637
638   stcox i.Hwv8_ca2_depress_group H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.
      H_wealthquintiles ///
639   i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity
640
641
642
643
644   * Coefficients instead of hazard ratios by specifing the option nohr
645
646   stcox i.Hwv8_ca2_depress_group, nohr
647
648   stcox i.Hwv8_ca2_depress_group H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.
      H_wealthquintiles ///
649   i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity, nohr
650
651
652
653
654   * Multivariable model development
655   * Likelihood-ratio tests
656
657
658
659   *install eststo
660   findit eststo
661
662
663   * ---- rx controlling for age and sex -----*
664   quietly: stcox H_age i.H_sex
665   eststo modelagesex
666
667   quietly: stcox H_age i.H_sex i.Hwv8_ca2_depress_group
668   eststo modelagesex_4group
```

```
669
670   lrtest modelagesex modelagesex_4group
671
672
673
674   * ---- rx controlling for sociodemographics -----*
675   quietly: stcox H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles
676   eststo modelsociodemo
677
678   quietly: stcox H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles i.
      Hwv8_ca2_depress_group
679   eststo modelsociodemo_4group
680
681   lrtest modelsociodemo modelsociodemo_4group
682
683
684   * ---- rx controlling for lifestyle/health indicators -----*
685   quietly: stcox i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity
686   eststo modelcardiovascular
687
688   quietly: stcox i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity i.Hwv8_ca2_depress_group
689   eststo modelcardiovascular_4group
690
691   lrtest modelcardiovascular modelcardiovascular_4group
692
693
694
695
696   * side-by-side comparison of models
697
698
699   quietly: stcox i.Hwv8_ca2_depress_group
700   eststo model1
701
702
703   quietly: stcox H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles i.
      Hwv8_ca2_depress_group
704   eststo model2
705
706
707   quietly: stcox H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
708   i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity i.Hwv8_ca2_depress_group
709   eststo model3
710
711
712
713   * Display Betas and Summary Statistics
714   estout model1 model2 model3, stats(n chi2 bic, star(chi2)) prehead("Betas")
715
716   /* Key Interpretattion
717   Chi2 = Value of LR test comparing the model fit (“full”) to intercept only (“reduced”)
718   bic = Schwarz’ Bayesian Information Criterion = It is a function of the log-likelihood.
719   Smaller values indicate a better fit.
720   */
721
722   * Display Hazard Ratios and Model Fit Statistics. Option eform produces hazard ratios
723   estout model1 model2 model3, eform stats(n chi2 bic, star(chi2)) prehead("Hazard Ratios")
724
725
726
727
728   * Postestimation tools for stcox
729
730   /* Test of proportional hazards
731
732   If the tests in the table are not significance (p-values over 0.05)
733    then we can not reject proportionality and we assume
734    that we do not have a violation of the proportional assumption.
```

```
*/

estat phtest, detail


/* Proportionality Assumption - method 1
We will check proportionality by including time-dependent covariates in the model
by using the tvc and the texp options in the stcox command.
Time dependent covariates are interactions of the predictors and time.
In this analysis we choose to use the interactions with log(time)
because this is the most common function of time used in time-dependent covariates
but any function of time could be used.
If a time-dependent covariate is significant this indicates
a violation of the proportionality assumption for that specific predictor.
The conclusion is that all of the time-dependent variables are not significant
either collectively or individually thus supporting the assumption of proportional hazard.
*/



stcox i.Hwv8_ca4_depress_group H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.
H_wealthquintiles ///
i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity, nohr ///
tvc(Hwv8_ca2_depress_group H_age H_sex H_eduaction H_maritalstatus_4cat H_wealthquintiles ///
H_smoking_3cat H_alcohol_status H_cvd_comorbidity) texp(ln(H_time_of_event_dementia))



/* Proportionality Assumption - method 2
by using the Schoenfeld and scaled Schoenfeld residuals
In the stphtest command we test the proportionality of the model as a whole
and by using the detail option we get a test of proportionality for each predictor.
By using the plot option we can also obtain a graph of the scaled Schoenfeld assumption.
If the tests in the table are not significance (p-values over 0.05)
then we can not reject proportionality and we assume
that we do not have a violation of the proportional assumption.
The stphplot command uses log-log plots to test proportionality
and if the lines  in these plots are parallel then we have further indication
that the predictors do not violate the proportionality assumption.
*/

quietly stcox Hwv8_ca2_depress_group H_age H_sex H_eduaction H_maritalstatus_4cat
H_wealthquintiles ///
H_smoking_3cat H_alcohol_status H_cvd_comorbidity, schoenfeld(sch*) scaledsch(sca*)
stphtest, detail
stphtest, plot(Hwv8_ca2_depress_group) msym(oh)
stphtest, plot(H_age) msym(oh)
stphtest, plot(H_sex) msym(oh)
stphtest, plot(H_eduaction) msym(oh)
stphtest, plot(H_maritalstatus_4cat) msym(oh)
stphtest, plot(H_wealthquintiles) msym(oh)
stphtest, plot(H_smoking_3cat) msym(oh)
stphtest, plot(H_alcohol_status) msym(oh)
stphtest, plot(H_cvd_comorbidity) msym(oh)




stphplot, by(Hwv8_ca2_depress_group) plot1(msym(oh)) plot2(msym(th))
stphplot, by(H_age) plot1(msym(oh)) plot2(msym(th))
stphplot, by(H_sex) plot1(msym(oh)) plot2(msym(th))
stphplot, by(H_eduaction) plot1(msym(oh)) plot2(msym(th))
stphplot, by(H_maritalstatus_4cat) plot1(msym(oh)) plot2(msym(th))
stphplot, by(H_wealthquintiles) plot1(msym(oh)) plot2(msym(th))
stphplot, by(H_smoking_3cat) plot1(msym(oh)) plot2(msym(th))
stphplot, by(H_alcohol_status) plot1(msym(oh)) plot2(msym(th))
stphplot, by(H_cvd_comorbidity) plot1(msym(oh)) plot2(msym(th))
```

```
801
802
803
804
805   * Assessment of PH Assumption: adjust for age and sex
806   stphplot, by(Hwv8_ca2_depress_group) adjust(H_age H_sex) nolntime plot1opts(symbol(none) color(
      black) lpattern(dash)) ///
807   plot2opts(symbol(none) color(green)) plot3opts(symbol(none) color(red)) ///
808   title("Assessment of PH Assumption") subtitle(" Predictor is Hwv8_ca4_depress_group") xtitle(
      "months")
809
810
811
812   * Assessment of PH Assumption: adjust for model 2
813   stphplot, by(Hwv8_ca2_depress_group) adjust(H_age H_sex H_eduaction H_maritalstatus_4cat
      H_wealthquintiles) ///
814   nolntime plot1opts(symbol(none) color(black) lpattern(dash)) ///
815   plot2opts(symbol(none) color(green)) plot3opts(symbol(none) color(red)) ///
816   title("Assessment of PH Assumption") subtitle(" Predictor is Hwv8_ca4_depress_group") xtitle(
      "months")
817
818
819
820   * Assessment of PH Assumption: adjust for model 3
821   stphplot, by(Hwv8_ca2_depress_group) adjust(H_age H_sex H_eduaction H_maritalstatus_4cat
      H_wealthquintiles ///
822   H_smoking_3cat H_alcohol_status H_cvd_comorbidity) ///
823   nolntime plot1opts(symbol(none) color(black) lpattern(dash)) ///
824   plot2opts(symbol(none) color(green)) plot3opts(symbol(none) color(red)) ///
825   title("Assessment of PH Assumption") subtitle(" Predictor is Hwv8_ca4_depress_group") xtitle(
      "months")
826
827
828
829
830
831
832   /* Test of overall goodness of fit
833   Goodness of fit of the final model
834   2 methods:
835    - by using the commnad stcoxgof (good fit = non sig p-value)
836    - by using the Cox-Snell residuals
837       to create the Nelson-Aalen cumulative hazard function
838       If the hazard function follows the 45 degree line then we know that it approximately
839       has an exponential distribution with a hazard rate of one and that the model fits the data
      well.
840       If the model fits the data, the plot of the cumulative hazard versus cs
841       should approximate a straight line with slope 1.
842   */
843
844
845   * by using the commnad stcoxgof
846
847   * install stcoxgof
848   findit stcoxgof
849
850
851   stcox Hwv8_ca2_depress_group H_age H_sex H_eduaction H_maritalstatus_4cat H_wealthquintiles ///
852   H_smoking_3cat H_alcohol_status H_cvd_comorbidity, mgale(mgale)
853
854
855   stcoxgof
856
857
858
859   * by using the Cox-Snell residuals
860
861   quietly stcox Hwv8_ca2_depress_group H_age H_sex H_eduaction H_maritalstatus_4cat
```

```
      H_wealthquintiles ///
862   H_smoking_3cat H_alcohol_status H_cvd_comorbidity
863   predict cs, csnell
864
865   * or
866
867   quietly stcox Hwv8_ca2_depress_group
868   predict cs, csnell
869
870
871   stset cs, failure(Hwv9to14_dementia_event)
872   sts generate km = s
873   generate H = -ln(km)
874   line H cs cs, sort ytitle("") clstyle(. refline)
875
876
877
878
879
880   /* Cox PH regression model for independent depressive symptoms and CA exposure variable
881
882   Hwv8_depressive_symptoms
883   Hwv8_crp
884   Hwv8_hdl
885   Hwv8_obesity_waist
886   Hwv8_systolic_bp
887   Hwv8_diastolic_bp
888   Hwv8_diabetes_reportevr
889   Hwv8_HbA1c
890   Hwv8_cardio_number
891   Hwv8_cardio2
892   Hwv8_cardio3
893   Hwv8_cardio4
894
895   */
896
897
898
899   stset H_time_of_event_dementia, failure (Hwv9to14_dementia_event==1) id(RAHHIDPN)
900
901
902
903
904   * Unadjusted model 1
905
906
907   stcox i.Hwv8_depressive_symptoms
908
909
910   * Adjusted models - multivariable Cox model
911   * controlling for covariates
912
913   * model 2: model 1 + adjust for socio-demographics: age sex education marital status and wealth
914
915   stcox i.Hwv8_depressive_symptoms H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.
      H_wealthquintiles
916
917
918   * model 3: model 2 + adjust for lifestyle/health indicators
919
920   stcox i.Hwv8_depressive_symptoms H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.
      H_wealthquintiles ///
921   i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity
922
923
924
925   * repeat for each independent variable from the list above
926
```

```
927
928
929
930
931   ** ---------------------- **
932
933
934
935
936
937   /* MULTIPLE IMPUTATION (MI)
938
939   To handle with missing values of covariates
940
941
942   useful sources for MI and MICE:
943
944   https://stats.idre.ucla.edu/stata/seminars/mi_in_stata_pt1_new/
945   https://www.stata.com/manuals/mi.pdf - see page 139
946   https://www.stata.com/meeting/switzerland16/slides/medeiros-switzerland16.pdf
947   https://www.youtube.com/watch?v=i6SOlq0mjuc&ab_channel=StataCorpLLC
948   https://dss.princeton.edu/training/MIStata.pdf
949
950
951
952   Preparing to conduct MI
953   1. examine the number and proportion of missing values among the variables of interest
954       use the mdesc command
955   2. examine missing data patterns
956       use commands mi set and mi misstable patterns
957   3. identify potential auxiliary variables
958
959
960   Run MI using chained equations (MICE)
961   using the commands
962   1. how (in what style) to store the imputations
963    mi set wide
964   2. which variables will be imputed
965    mi register imputed
966   3. optionally, which variables will not be imputed
967    mi register regular
968   4. what imputation method is implemented to impute each of var - MICE
969    mi impute chained
970
971   */
972
973
974
975
976
977   /*
978
979   1. examining missing values
980       install packages:
981       * install mdesc
982       * install tabmiss
983       * insatll dm31
984       * insall mvpatterna
985
986   */
987
988   search mdesc
989   search rmiss2
990   search mvpatterns
991
992
993
994
```

```
995
996
997   * examining number of missing values vs non-missing in each variable
998
999   mdesc H_age H_sex H_eduaction H_maritalstatus_4cat H_wealthquintiles ///
1000  H_smoking_3cat H_physicalactivity H_alcohol_status H_cvd_comorbidity Hwv8_memory
1001
1002
1003
1004
1005
1006  * examining missing data patterns
1007
1008  mi set wide
1009
1010  mi misstable summarize H_age H_sex H_eduaction H_maritalstatus_4cat H_wealthquintiles ///
1011  H_smoking_3cat H_physicalactivity H_alcohol_status H_cvd_comorbidity
1012
1013
1014  mi misstable patterns H_age H_sex H_eduaction H_maritalstatus_4cat H_wealthquintiles ///
1015  H_smoking_3cat H_physicalactivity H_alcohol_status H_cvd_comorbidity
1016
1017
1018
1019  /*
1020   identifying potential auxiliary var
1021  * Auxiliary variables are either correlated with a missing variable(s)
1022  (the recommendation is r > 0.4) or are believed to be associated with missingness
1023  - a priori knowledge of var that would make good auxiliary var
1024  - identify potential candidates by examining associations between missing var and other var in
      the dataset
1025      running correlation using the command: pwcorr v1 v2 v3, obs
1026      the recommnedation for good correlation is r > 0.4
1027
1028
1029  Missing var to be imputed:
1030
1031      H_smoking_3cat H_physicalactivity H_alcohol_status
1032
1033
1034  Potential auxiliary var:
1035  DV:  Hwv9to14_dementia_event
1036  IV:  Hwv8_depressive_symptoms Hwv8_crp Hwv8_hdl Hwv8_obesity_waist Hwv8_systolic_bp
      Hwv8_diastolic_bp Hwv8_diabetes_reportevr Hwv8_HbA1c
1037  other var:
1038      H_age H_sex H_eduaction H_maritalstatus_4cat H_wealthquintiles H_cvd_comorbidity
1039
1040  */
1041
1042
1043  * correlation
1044
1045  pwcorr H_smoking_3cat H_physicalactivity H_alcohol_status ///
1046      Hwv9to14_dementia_event ///
1047      Hwv8_depressive_symptoms Hwv8_crp Hwv8_hdl Hwv8_obesity_waist Hwv8_systolic_bp
      Hwv8_diastolic_bp Hwv8_diabetes_reportevr Hwv8_HbA1c ///
1048      H_age H_sex H_eduaction H_maritalstatus_4cat H_wealthquintiles H_cvd_comorbidity, obs
1049
1050
1051  /* The correlation showed that all the following var are good auxiliary:
1052  Hwv9to14_dementia_event Hwv8_depressive_symptoms Hwv8_diabetes_reportevr Hwv8_HbA1c H_age H_sex
      H_eduaction H_maritalstatus_4cat H_wealthquintiles H_cvd_comorbidity
1053   A good auxiliary does not have to be correlated with every variable to be useful
1054   And it's not problematic if it has missing info of it's own
1055  */
1056
1057
1058
```

```
/*
MI by chained equations (MICE)
    see: https://stats.idre.ucla.edu/stata/seminars/mi_in_stata_pt1_new/

MICE is known as the fully conditional specification or sequential generalized regression
does not assume a joint MVN distribution
but instead uses a separate conditional distribution for each imputed variable.

The multivariate normal (MVN) model - mi imputed mvn -
assumes multivariate normality of all var

The multivariate imputation by chained equations (MICE) - mi imputed chained -
offers flexibility in how each var is modeled

mi impute chained allows to specify models for a
variety of variable types, including
continuous, binary, ordinal, nominal, truncated, and count variables


The MICE distributions available in Stata are:
binary, ordered and multinomial logistic regression for categorical variables,
linear regression and predictive mean matching (PMM)* for continuous variables,
and Poisson and negative binomial regression for count variables.



IMPUTATION PHASES

1. mi set wide
    style to store imputations

2. mi register imputed
    identifies which variables in the imputation model have missing information.

3. mi register regular (! optional)
    which variables will not be imputed

4. mi impute chained
    where the user specifies the imputation model to be used
    and the number of imputed datasets to be created.
    Example:
        mi impute chained (regress) bmi age (logit) female ///
        (mlogit) race = bpdiast i.region, add(20)

5. mi estimate
    is used as a prefix to the standard regress command.
    This executes the specified estimation model within each of the 20 imputed datasets
    to obtain 20 sets of coefficients and standard errors.
    Stata then combines these estimates to obtain one set of inferential statistics.
    In the output from mi estimate you will see some metrics: Imputation Diagnostics
    information for RVI (Relative Increase in Variance),
    FMI (Fraction of Missing Information),
    DF (Degrees of Freedom) ,
    RE (Relative Efficiency),
    and the between imputation and the within imputation variance estimates
    to examine how the standard errors (SEs) are calculated.


----------------

SELECTING MY IMPUTATION MODEL

- MICE -> mi impute chained

- var to be imputed:

    linear regression for continuous var (regress) -> none

```

```
1127
1128        logistic for the binary var (logit) -> none
1129
1130        multinomial logistic for our nominal categorical var (mlogit) ->
1131        H_smoking_3cat H_physicalactivity H_alcohol_status
1132
1133
1134
1135    - auxiliary var:
1136
1137        DV -> Hwv9to14_dementia_event
1138        IV -> Hwv8_depressive_symptoms Hwv8_diabetes_reportevr Hwv8_HbA1c
1139        other covariates -> H_age H_sex H_eduaction H_maritalstatus_4cat H_wealthquintiles
    H_cvd_comorbidity
1140
1141
1142
1143    - imputation numbers (m) -> 10
1144
1145        ELSA data were imputed 10 numbers
1146
1147        White et al. (2010) recommendation: use the rule that m should equal the percentage of
    incomplete cases
1148
1149
1150    - rseed (53421) for reproducability reasons
1151
1152
1153    - (! OPTIONAL) advance impute options -> force
1154
1155        proceed with imputation, even when missing imputed values (e.g. auxiliary have missing data)
    are encountered
1156
1157    - impute options -> savetrace (trace1)
1158
1159        specifies Stata to save the means and standard deviations of imputed values from each
    iteration to a Stata dataset named “trace1
1160    */
1161
1162
1163    mi set wide
1164
1165
1166    mi register imputed H_smoking_3cat H_physicalactivity H_alcohol_status
1167
1168
1169
1170    mi impute chained (mlogit) H_smoking_3cat H_physicalactivity H_alcohol_status =
    Hwv8_depressive_symptoms Hwv8_diabetes_reportevr Hwv8_HbA1c H_age H_sex H_eduaction
    H_maritalstatus_4cat H_wealthquintiles H_cvd_comorbidity, add(10) rseed(53421) savetrace(trace1)
1171
1172
1173    * save imputed data
1174
1175
1176    * plot imputations
1177
1178    *it will open a file named trace1
1179    use trace1,clear
1180    describe
1181
1182
1183    reshape wide *mean *sd, i(iter) j(m)
1184    tsset iter
1185
1186
1187
1188    /*
```

```
The trace plot below graphs the predicted means value produced during the first imputation chain.
As before, the expectations is that the values would vary randomly to incorporate variation into the predicted values for read.
*/

tsline H_smoking_3cat_mean1, name(mice1,replace)legend(off) ytitle("Mean of smoking")
tsline H_physicalactivity_mean1, name(mice1,replace)legend(off) ytitle("Mean of physical activity")
tsline H_alcohol_status_mean1, name(mice1,replace)legend(off) ytitle("Mean of alcohol status")


/*

All imputation chains can also be graphed simultaneously to make sure that nothing unexpected occurred in a single chain.
Every chain is obtained using a different set of initial values and this should be unique.
Each colored line represents a different imputation.
So all 10 imputation chains are overlaid on top of one another.

*/


tsline H_alcohol_status_mean*, name(mice1,replace)legend(off) ytitle("Mean of alcohol")
tsline H_alcohol_status_sd*, name(mice2, replace) legend(off) ytitle("SD of alcohol")
graph combine mice1 mice2, xcommon cols(1) title(Trace plots of summaries of imputed values)

* repeat for each imputed var









* -------- COX PH REGRESSION MODEL IN IMPUTED DATASET --------- *


* Declare Data to be Survival Data by using mi

mi stset H_time_of_event_dementia, failure (Hwv9to14_dementia_event==1) id(RAHHIDPN)

* Run Cox regression analysis in imputed dataset by using "mi estimate:"



/*
Independent risk factors

Hwv8_depressive_symptoms
Hwv8_crp
Hwv8_hdl
Hwv8_obesity_waist
Hwv8_systolic_bp
Hwv8_diastolic_bp
Hwv8_diabetes_reportevr
Hwv8_HbA1c
Hwv8_cardio_number
Hwv8_cardio2

*/




* Depressive symptoms


* Unadjusted model - model 1 - single predictor
```

```
1255
1256  * Model 1 (default coefficents)
1257  mi estimate: stcox Hwv8_depressive_symptoms
1258
1259  * Model 1: define design var by using i.
1260  mi estimate: stcox i.Hwv8_depressive_symptoms
1261
1262
1263  * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1264
1265  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_depressive_symptoms
1266
1267
1268  * Adjusted models - multivariable Cox model
1269  * controlling for covariates
1270
1271  * model 2: model 1 + adjust for sociodemographics: age sex education marital status and wealth
1272
1273  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_depressive_symptoms ///
1274  H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles
1275
1276
1277  * model 3: model 2 + adjust for lifestyle/ health indicators
1278
1279  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_depressive_symptoms ///
1280  H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
1281  i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity
1282
1283
1284
1285  * repeat for each independent variable from the list above
1286
1287
1288
1289
1290
1291  /*
1292  Combined effects Cox regression models
1293
1294  Hwv8_crp_depress_group
1295  Hwv8_hdl_depress_group
1296  Hwv8_waist_depress_group
1297  Hwv8_sbp_depress_group
1298  Hwv8_dbp_depress_group
1299  Hwv8_diabet_depress_group
1300  Hwv8_hba1c_depress_group
1301  Hwv8_ca2_depress_group
1302
1303  */
1304
1305
1306  * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1307
1308  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group
1309
1310
1311  * Adjusted models - multivariable Cox model
1312  * controlling for covariates
1313
1314  * model 2: model 1 + adjust for sociodemographics: age sex education marital status and wealth
1315
1316  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group ///
1317  H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles
1318
1319
1320  * model 3: model 2 + adjust for lifestyle/ health indicators
1321
1322  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group ///
```

```
1323   H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
1324   i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity
1325
1326
1327
1328   * repeat for each variable from the list above
1329
1330
1331
1332
1333   *-------------------------*
1334
1335
1336   /*
1337
1338   *** SENSITIVITY ANALYSES ***
1339
1340
1341   1) interarction effect of gender and age_group
1342
1343   2) survival analysis stratified by age
1344   two age groups: young old <70 and old old >=70
1345
1346   3) depressive symptoms as continuous variable
1347   and >= 3 and >=4 cardiometabolic multimorbidity
1348
1349
1350   4) exclude participants with cvd
1351
1352   5) Complete data
1353
1354   6) survival analysis limiting to 5 year follow-up period
1355
1356
1357
1358   Repeat on all independent and combined variables
1359
1360   Hwv8_depressive_symptoms
1361   Hwv8_crp
1362   Hwv8_hdl
1363   Hwv8_obesity_waist
1364   Hwv8_systolic_bp
1365   Hwv8_diastolic_bp
1366   Hwv8_diabetes_reportevr
1367   Hwv8_HbA1c
1368   Hwv8_cardio_number
1369   Hwv8_cardio2
1370
1371
1372   Hwv8_crp_depress_group
1373   Hwv8_hdl_depress_group
1374   Hwv8_waist_depress_group
1375   Hwv8_sbp_depress_group
1376   Hwv8_dbp_depress_group
1377   Hwv8_diabet_depress_group
1378   Hwv8_hba1c_depress_group
1379   Hwv8_ca2_depress_group
1380
1381
1382   */
1383
1384
1385
1386   * 1) Interaction effect
1387
1388   * sex*risk factor
1389
1390   mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio2 i.H_sex#i.Hwv8_cardio2
```

```
1391
1392  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio2 ///
1393  H_age i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
1394  i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity ///
1395   i.H_sex#i.Hwv8_cardio2
1396
1397
1398
1399  * age*risk factor
1400
1401  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio2 c.H_age#i.Hwv8_cardio2
1402
1403  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio2 ///
1404  H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
1405  i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity ///
1406   c.H_age#i.Hwv8_cardio2
1407
1408
1409
1410
1411  /* 2) Survival analysis stratified by age
1412
1413  generate age group variable
1414  Age groups: 1) young old (< 70) 2) old old (>= 70)
1415
1416  Kaplan Meier curves
1417  Cox regression models in imputed data
1418
1419  young old <70
1420  if H_age_group==1
1421
1422  old old >70
1423  if H_age_group==2
1424
1425
1426  */
1427
1428
1429
1430  gen H_age_group=1 if H_age < 70
1431  replace H_age_group=2 if H_age >=70 & ///
1432  !missing(H_age)
1433
1434  label var H_age_group "Age groups <70 young-old / 70 old-old"
1435  lab def age_group 1 "young old <70" 2 "old old >70"
1436  lab val H_age_group age_group
1437
1438  tab H_age_group
1439
1440
1441
1442
1443
1444  * COX PH REGRESSION MODEL IN IMPUTED DATASET
1445
1446  * Declare Data to be Survival Data by using mi
1447
1448  mi stset H_time_of_event_dementia, failure (Hwv9to14_dementia_event==1) id(RAHHIDPN)
1449
1450
1451
1452  * YOUNG OLD <70 Cox regression models
1453
1454  * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1455
1456  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group if H_age_group==1
1457
1458
```

```
1459
1460  * Model 3: model 2 + adjust for lifestyle/health indicators
1461
1462  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group ///
1463   i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
1464  i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity if H_age_group==1
1465
1466
1467  * OLD OLD >70 Cox regression models
1468
1469  * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1470
1471  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group if H_age_group==2
1472
1473
1474  * Model 3: model 2 + adjust for lifestyle/health indicators
1475
1476  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group ///
1477  i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
1478  i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity if H_age_group==2
1479
1480
1481
1482
1483
1484
1485  * 3) On depressive symptoms continuous variable
1486
1487
1488  * COX PH REGRESSION MODEL IN IMPUTED DATASET
1489
1490
1491  * Declare Data to be Survival Data by using mi
1492
1493  mi stset H_time_of_event_dementia, failure (Hwv9to14_dementia_event==1) id(RAHHIDPN)
1494
1495
1496  * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1497
1498  mi estimate, eform("Haz. Ratio"): stcox Hwv8_cesd_sumscore
1499
1500  * Adjusted models - multivariable Cox model
1501  * controlling for covariates
1502
1503  * Model 2: model 1 + adjust for socio-demographics: age sex education marital status and wealth
1504  mi estimate, eform("Haz. Ratio"): stcox Hwv8_cesd_sumscore ///
1505   H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles
1506
1507
1508
1509  * Model 3: model 2 + adjust for lifestyle/health indicators
1510
1511  mi estimate, eform("Haz. Ratio"): stcox Hwv8_cesd_sumscore ///
1512  H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
1513  i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity
1514
1515
1516
1517
1518
1519  * Cardiometabolic multimorbidty >= 3
1520
1521
1522
1523  * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1524
1525  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio3
1526
```

```
* Adjusted models - multivariable Cox model
* controlling for covariates

* Model 2: model 1 + adjust for socio-demographics: age sex education marital status and wealth
mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio3 ///
 H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles



* Model 3: model 2 + adjust for lifestyle/health indicators

mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio3 ///
H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity




* Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")

mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca3_depress_group

* Adjusted models - multivariable Cox model
* controlling for covariates

* Model 2: model 1 + adjust for socio-demographics: age sex education marital status and wealth
mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca3_depress_group ///
 H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles



* Model 3: model 2 + adjust for lifestyle/health indicators

mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca3_depress_group ///
H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity




* Cardiometabolic multimorbidty >=4



* Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")

mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio4

* Adjusted models - multivariable Cox model
* controlling for covariates

* Model 2: model 1 + adjust for socio-demographics: age sex education marital status and wealth
mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio4 ///
 H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles



* Model 3: model 2 + adjust for lifestyle/health indicators

mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_cardio4 ///
H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity




* Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")

```

```
mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca4_depress_group

* Adjusted models - multivariable Cox model
* controlling for covariates

* Model 2: model 1 + adjust for socio-demographics: age sex education marital status and wealth
mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca4_depress_group ///
 H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles



* Model 3: model 2 + adjust for lifestyle/health indicators

mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca4_depress_group ///
H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity









/*

4) exclude participants with cvd

use the command if H_cvd_comorbidity==0

*/



* without cvd

* Declare Data to be Survival Data by using mi

mi stset H_time_of_event_dementia, failure (Hwv9to14_dementia_event==1) id(RAHHIDPN)

* Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")

mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group if H_cvd_comorbidity==0



* Model 3: model 2 + adjust for lifestyle/health indicators

mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group ///
H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
i.H_smoking_3cat i.H_alcohol_status if H_cvd_comorbidity==0







* 5) On complete cases (see above)



/*
6) survival analysis limiting to 5 year follow-up period

hrs follow-up wave 9-12

*/
```

```
1663
1664
1665
1666  merge 1:m RAHHIDPN using "S:\Research\pkstudies\Study3_cardio_lca\HRS\hrs_lca data sensitivity
      9to12followup.dta"
1667
1668
1669
1670  * Declare Data to be Survival Data by using mi
1671
1672  mi stset Hwv9to12_time_of_event_dementia, failure (Hwv9to12_dementia_event==1) id(RAHHIDPN)
1673
1674
1675
1676  * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1677
1678  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group
1679
1680
1681  * Adjusted models - multivariable Cox model
1682  * controlling for covariates
1683
1684  * Model 2: model 1 + adjust for socio-demographics: age sex education marital status and wealth
1685  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group ///
1686   H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles
1687
1688
1689
1690  * Model 3: model 2 + adjust for lifestyle/health indicators
1691
1692  mi estimate, eform("Haz. Ratio"): stcox i.Hwv8_ca2_depress_group ///
1693  H_age i.H_sex i.H_eduaction i.H_maritalstatus_4cat i.H_wealthquintiles ///
1694  i.H_smoking_3cat i.H_alcohol_status i.H_cvd_comorbidity
1695
1696
1697
1698
1699
1700
1701  * ------------------------------ *
1702
1703
1704
1705
1706
1707
```