```
/*

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: ELSA
baseline: wave 2 (2004) follow-up waves 3-9 (2006-2018)


TIMELINE

DEPRESSIVE SYMPTOMS AND CARDIOMETABOLIC RISK FACTORS: WV2 (BASELINE)
DEMENTIA INCIDENCE: W3 - WV9 (7 TIME POINTS)
COVARIATES ADJUSTMENT FOR HR MODELS: WV2

*/



* KEEP NECESSARY VARIABLES

keep idauniq w2wtnur w2wtbld ///
E_sex E_age E_eduaction_yrs E_eduaction E_maritalstatus_3cat E_maritalstatus_4cat ///
E_wealthquintiles E_smoking_3cat E_physicalactivity E_alcohol_freq E_alcohol_status ///
E_cvd_comorbidity E_cognitive_index E_memory_wordrecall Ewv6_memory_wordrecall ///
Ewv2_loneliness_quintiles ///
Ewv2_cesd_score Ewv2_depressive_symptoms ///
Ewv3_cesd_sumscore_rand Ewv3_depressive_symptoms ///
Ewv4_cesd_sumscore_rand Ewv4_depressive_symptoms ///
Ewv5_cesd_sumscore_rand Ewv5_depressive_symptoms ///
Ewv6_cesd_sumscore_rand Ewv6_depressive_symptoms ///
Ewv7_cesd_sumscore_rand Ewv7_depressive_symptoms ///
Ewv8_cesd_sumscore Ewv8_depressive_symptoms ///
Ewv9_cesd_sumscore Ewv9_depressive_symptoms ///
Ewv2_crp_level Ewv2_crp Ewv2_fibrinogen_level Ewv2_fibrinogen ///
Ewv2_hdl_level Ewv2_male_hdl Ewv2_female_hdl ///
Ewv2_meds_hdl Ewv2_cholesterol_evr Ewv2_hdl_sum Ewv2_hdl_cholesterol ///
Ewv2_waist Ewv2_malewaist_ao Ewv2_femalewaist_ao Ewv2_obesity_waist_sum Ewv2_obesity_waist ///
Ewv2_bmi_score Ewv2_obesity_bmi Ewv2_waist_bmi_sum Ewv2_obesity ///
Ewv2_tg_level Ewv2_tg ///
Ewv2_systolic_mean Ewv2_diastolic_mean Ewv2_systolic_bp Ewv2_diastolic_bp ///
Ewv2_meds_bp Ewv2_bp_reportevr Ewv2_bp_before Ewv2_bp_diagnosed_sum Ewv2_bp_diagnosed Ewv2_bp_sum Ewv2_bp ///
Ewv2_diabetes_evr Ewv2_diabetes_before Ewv2_diabetes_diagnosed_sum Ewv2_diabetes_diagnosed ///
Ewv2_glucose_level Ewv2_glucose Ewv2_HbA1c_level Ewv2_HbA1c ///
Ewv2_meds1_diabetes Ewv2_meds2_diabetes Ewv2_insulin_diabetes ///
Ewv2_diabetes_anymeds_sum Ewv2_diabetes_anymeds ///
Ewv2_diabetes_glucose_sum Ewv2_glycemia ///
Ewv2_ao_depress_sum Ewv2_Nao_Ndepress Ewv2_Nao_Ydepress Ewv2_Yao_Ndepress Ewv2_Yao_Ydepress ///
Ewv2_ao_depress_group Ewv2_waist_depress_sum Ewv2_Nwaist_Ndepress Ewv2_Nwaist_Ydepress ///
Ewv2_Ywaist_Ndepress Ewv2_Ywaist_Ydepress Ewv2_waist_depress_group Ewv2_glycemia_depress_sum ///
Ewv2_Nglycemia_Ndepress Ewv2_Nglycemia_Ydepress Ewv2_Yglycemia_Ndepress Ewv2_Yglycemia_Ydepress ///
Ewv2_glycemia_depress_group Ewv2_diabet_depress_sum Ewv2_Ndiabet_Ndepress Ewv2_Ndiabet_Ydepress ///
Ewv2_Ydiabet_Ndepress Ewv2_Ydiabet_Ydepress Ewv2_diabet_depress_group Ewv2_hba1c_depress_sum ///
Ewv2_Nhba1c_Ndepress Ewv2_Nhba1c_Ydepress Ewv2_Yhba1c_Ndepress Ewv2_Yhba1c_Ydepress ///
Ewv2_hba1c_depress_group Ewv2_hdl_depress_sum Ewv2_Nhdl_Ndepress Ewv2_Nhdl_Ydepress ///
Ewv2_Yhdl_Ndepress Ewv2_Yhdl_Ydepress Ewv2_hdl_depress_group Ewv2_bp_depress_sum Ewv2_Nbp_Ndepres ///
Ewv2_Nbp_Ydepress Ewv2_Ybp_Ndepress Ewv2_Ybp_Ydepress Ewv2_bp_depress_group Ewv2_sbp_depress_sum ///
Ewv2_Nsbp_Ndepress Ewv2_Nsbp_Ydepress Ewv2_Ysbp_Ndepress Ewv2_Ysbp_Ydepress Ewv2_sbp_depress_group ///
Ewv2_dbp_depress_sum Ewv2_Ndbp_Ndepress Ewv2_Ndbp_Ydepress Ewv2_Ydbp_Ndepress Ewv2_Ydbp_Ydepress ///
```

```
Ewv2_dbp_depress_group Ewv2_crp_depress_sum ///
Ewv2_Ncrp_Ndepress Ewv2_Ncrp_Ydepress ///
Ewv2_Ycrp_Ndepress Ewv2_Ycrp_Ydepress ///
Ewv2_crp_depress_group Ewv2_tg_depress_sum ///
Ewv2_Ntg_Ndepress Ewv2_Ntg_Ydepress Ewv2_Ytg_Ndepress Ewv2_Ytg_Ydepress ///
Ewv2_tg_depress_group Ewv2_fibr_depress_sum Ewv2_Nfibr_Ndepress ///
Ewv2_Nfibr_Ydepress Ewv2_Yfibr_Ndepress ///
Ewv2_Yfibr_Ydepress Ewv2_fibr_depress_group ///
Ewv2_cardio3 Ewv2_ca3_depress_sum Ewv2_Nca3_Ndepress Ewv2_Nca3_Ydepress ///
Ewv2_Yca3_Ndepress Ewv2_Yca3_Ydepress Ewv2_ca3_depress_group ///
Ewv2_cardio4 Ewv2_ca4_depress_sum Ewv2_Nca4_Ndepress Ewv2_Nca4_Ydepress ///
Ewv2_Yca4_Ndepress Ewv2_Yca4_Ydepress Ewv2_ca4_depress_group ///
Ewv2_cardio_number_sum Ewv2_cardio_number ///
Ewv2_cardio2 Ewv2_ca2_depress_sum ///
Ewv2_Nca2_Ndepress Ewv2_Nca2_Ydepress ///
Ewv2_Yca2_Ndepress Ewv2_Yca2_Ydepress Ewv2_ca2_depress_group ///
Ewv2_anydementia_iqcode_report ///
Ewv3_anydementia_iqcode_report ///
Ewv4_anydementia_iqcode_report Ewv6to9_dementia_event ///
Ewv5_anydementia_iqcode_report Ewv6_anydementia_iqcode_report Ewv7_anydementia_iqcode_report ///
Ewv8_anydementia_iqcode_report Ewv9_anydementia_iqcode_report ///
Ewv2_interview_date Ewv3_interview_date Ewv4_interview_date ///
Ewv5_interview_date Ewv6_interview_date Ewv7_interview_date ///
Ewv8_interview_date Ewv9_interview_date ///
Ewv3to9_dementia_sum Ewv3to9_dementia_sum_no_iqcode ///
Ewv3to9_dementia_event Ewv3to9_dementia_event_no_iqcode ///
Ewv3to9_dementia_report_or_lasti Ewv3to9_dementia_report_free_dat ///
Ewv3to9_newdementia_or_lastinter Ewv3to9_dementia_free_date E_time_dementia_months ///
E_time_dementia_report_months_no E_time_dementia_midpoint ///
E_time_dementia_midpoint_final E_time_event_dementia E_time_dementia_report_midpoint_ ///
E_time_dementia_midpoint_no_iqco E_time_event_dementia_report_no_ ///
Ewv6to9_dementia_free_date Ewv6to9_newdementia_or_lastinter ///
Ewv6to9_time_dementia_months Ewv6to9_time_dementia_midpoint ///
Ewv6to9_time_dementia_midpoint_f Ewv6to9_time_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 idauniq using
"S:\Research\pkstudies\Study1_biopsych_risk\ELSA\elsa_to_merge_physicalact.dta"

keep if _merge==3

drop _merge
```

```
129
130
131   /*
132
133   EXPOSURE VARIABLES
134
135
136   Binary variables of depressive symptoms and cardiometabolic markers measured at wave 2
137
138   Depression: Ewv2_depressive_symptoms
139
140   CRP: Ewv2_crp
141
142   HDL cholesterol: Ewv2_hdl_cholesterol
143
144   Obesity by waist cir: Ewv2_obesity_waist
145
146   systolic Blood pressure: Ewv2_systolic_bp
147
148   diastolic Blood pressure: Ewv2_diastolic_bp
149
150   Diabetes: Ewv2_diabetes_diagnosed
151
152   HbA1c: Ewv2_HbA1c
153
154   CA number (categ 0,1,2,3,4+): Ewv2_cardio_number
155
156   CA mutlimorbidity >= 2 CA conditions: Ewv2_cardio2
157
158   Grouping of Dep-CA: Ewv2_waist_depress_group Ewv2_diabet_depress_group Ewv2_hba1c_depress_group
      Ewv2_hdl_depress_group Ewv2_sbp_depress_group Ewv2_dbp_depress_group Ewv2_crp_depress_group
      Ewv2_ca3_depress_group Ewv2_ca4_depress_group Ewv2_ca2_depress_group
159
160
161   OUTCOME VARIABLES
162
163   Dementia event: Ewv3to9_dementia_event
164
165   Time-to-event: E_time_event_dementia
166
167
168   */
169
170
171
172
173
174
175   *** Descriptive stats of var of interest
176
177
178   tabulate Ewv2_depressive_symptoms
179   summarize Ewv2_depressive_symptoms
180
181   misstable summarize Ewv2_depressive_symptoms
182   misstable patterns Ewv2_depressive_symptoms
183
184   tabulate Ewv2_crp
185   summarize Ewv2_crp
186
187   misstable summarize Ewv2_crp
188   misstable patterns Ewv2_crp
189
190   tabulate Ewv2_hdl_cholesterol
191   summarize Ewv2_hdl_cholesterol
192
193   misstable summarize Ewv2_hdl_cholesterol
194   misstable patterns Ewv2_hdl_cholesterol
```

```
195
196  tabulate Ewv2_obesity_waist
197  summarize Ewv2_obesity_waist
198
199  misstable summarize Ewv2_obesity_waist
200  misstable patterns Ewv2_obesity_waist
201
202  tabulate Ewv2_systolic_bp
203  summarize Ewv2_systolic_bp
204
205  misstable summarize Ewv2_systolic_bp
206  misstable patterns Ewv2_systolic_bp
207
208
209  tabulate Ewv2_diastolic_bp
210  summarize Ewv2_diastolic_bp
211
212  misstable summarize Ewv2_diastolic_bp
213  misstable patterns Ewv2_diastolic_bp
214
215
216  tabulate Ewv2_diabetes_diagnosed
217  summarize Ewv2_diabetes_diagnosed
218
219  misstable summarize Ewv2_diabetes_diagnosed
220  misstable patterns Ewv2_diabetes_diagnosed
221
222
223  tabulate Ewv2_HbA1c
224  summarize Ewv2_HbA1c
225
226  misstable summarize Ewv2_HbA1c
227  misstable patterns Ewv2_HbA1c
228
229
230  tabulate Ewv2_anydementia_iqcode_report
231  summarize Ewv2_anydementia_iqcode_report
232
233  misstable summarize Ewv2_anydementia_iqcode_report
234  misstable patterns Ewv2_anydementia_iqcode_report
235
236
237
238
239
240
241  *** CLEANING DATA
242
243
244  * 1. drop dementia cases and missing data at baseline
245
246  drop if Ewv2_anydementia_iqcode_report==1
247  * (50 observations deleted)
248
249  drop if Ewv2_anydementia_iqcode_report== .
250  * (0 observations deleted)
251
252
253  * 2. drop missing values of depresssive symptoms and cardiometabolic markers and invalid ca cases
254
255  drop if Ewv2_depressive_symptoms== .
256  * (73 observations deleted)
257
258  drop if Ewv2_crp_level > 100 & Ewv2_crp_level < 300
259  * (9 observations deleted)
260
261  drop if Ewv2_crp== .
262  * (1,733 observations deleted)
```

```
263
264  drop if Ewv2_hdl_cholesterol== .
265  * (6 observations deleted)
266
267  drop if Ewv2_obesity_waist== .
268  * (131 observations deleted)
269
270  drop if Ewv2_systolic_bp== .
271  * (650 observations deleted)
272
273  drop if Ewv2_diastolic_bp== .
274  * (0 observations deleted)
275
276  drop if Ewv2_diabetes_diagnosed== .
277  * (0 observations deleted)
278
279  drop if Ewv2_HbA1c== .
280  * (102 observations deleted)
281
282
283
284  * 3. drop obs with no records on dementia at any wave from 3-9 follow-ups
285
286
287  search mdesc
288  search rmiss2
289  search mvpatterns
290
291  * see number of missing values vs non-missing in each variable
292  mdesc Ewv3_anydementia_iqcode_report Ewv4_anydementia_iqcode_report ///
293  Ewv5_anydementia_iqcode_report Ewv6_anydementia_iqcode_report ///
294  Ewv7_anydementia_iqcode_report ///
295  Ewv8_anydementia_iqcode_report Ewv9_anydementia_iqcode_report
296
297
298
299  /* number of missing values per observation
300  * the code below creates a variable called nmisfollowup that gives the number of missing values
301  for each observation in the variables of interest */
302  egen nmisfollowup_dementia_wv3to9=rmiss2(Ewv3_anydementia_iqcode_report ///
303  Ewv4_anydementia_iqcode_report Ewv5_anydementia_iqcode_report ///
304  Ewv6_anydementia_iqcode_report Ewv7_anydementia_iqcode_report ///
305  Ewv8_anydementia_iqcode_report Ewv9_anydementia_iqcode_report)
306
307  tab nmisfollowup_dementia_wv3to9
308
309  * drop observations "nmisfollowup_dementia_wv3to9" > 6 (those with 7 missing data = no records at
     any wave)
310  drop if nmisfollowup_dementia_wv3to9>6
311  *(440 observations deleted)
312
313
314  * ANALYTIC SAMPLE -> 4472
315
316
317  /*
318  ---- DESCRIPTIVE STATISTICS ----
319
320  General characteristics of participnats stratified for study inclusion
321
322  General characteristics of participants stratified for dementia occurence
323
324  CHI-SQUARE (chi2) for categorical var (crosstabulation)
325      Frequency tables -> two-way tables
326          using the command tabulate, chi2
327          reporting observations, column percentage (N, %) and p-value of Pearson's r
328
329  one-way ANOVA for continuous var
```

```
330       check box plot
331       using the command oneway
332       reporting mean, sd (summary tables) and p-value of F
333   */
334
335
336   * General characteristics of ELSA participants at baseline
337
338   * Socio-demographics
339   sum E_age
340   ta E_sex
341   ta E_eduaction
342   ta E_maritalstatus_4cat
343   ta E_wealthquintiles
344   * Cardiometabolic risk factors
345   ta Ewv2_crp
346   ta Ewv2_hdl_cholesterol
347   ta Ewv2_obesity_waist
348   ta Ewv2_systolic_bp
349   ta Ewv2_diastolic_bp
350   ta Ewv2_diabetes_diagnosed
351   ta Ewv2_HbA1c
352   ta Ewv2_cardio2
353   * Lifestyle and health indicators
354   ta E_smoking_3cat
355   ta E_physicalactivity
356   ta E_alcohol_status
357   ta E_cvd_comorbidity
358   * Depressive symptoms (categ)
359   ta Ewv2_depressive_symptoms
360   * Memory score
361   sum E_memory_wordrecall
362
363
364   * General characteristics of ELSA participants stratified for dementia occurence
365
366   * Socio-demographics
367   ttest E_age, by(Ewv3to9_dementia_event)
368   ta E_sex Ewv3to9_dementia_event, chi2 column row
369   ta E_eduaction Ewv3to9_dementia_event, chi2 column row
370   ta E_maritalstatus_4cat Ewv3to9_dementia_event, chi2 column row
371   ta E_wealthquintiles Ewv3to9_dementia_event, chi2 column row
372   * Cardiometabolic risk factors
373   ta Ewv2_crp Ewv3to9_dementia_event, chi2 column row
374   ta Ewv2_hdl_cholesterol Ewv3to9_dementia_event, chi2 column row
375   ta Ewv2_obesity_waist Ewv3to9_dementia_event, chi2 column row
376   ta Ewv2_systolic_bp Ewv3to9_dementia_event, chi2 column row
377   ta Ewv2_diastolic_bp Ewv3to9_dementia_event, chi2 column row
378   ta Ewv2_diabetes_diagnosed Ewv3to9_dementia_event, chi2 column row
379   ta Ewv2_HbA1c Ewv3to9_dementia_event, chi2 column row
380   ta Ewv2_cardio2 Ewv3to9_dementia_event, chi2 column row
381   * Lifestyle and health indicators
382   ta E_smoking_3cat Ewv3to9_dementia_event, chi2 column row
383   ta E_physicalactivity Ewv3to9_dementia_event, chi2 column row
384   ta E_alcohol_status Ewv3to9_dementia_event, chi2 column row
385   ta E_cvd_comorbidity Ewv3to9_dementia_event, chi2 column row
386   * Depressive symptoms
387   ta Ewv2_depressive_symptoms Ewv3to9_dementia_event, chi2 column row
388   * Memory score
389   ttest E_memory_wordrecall, by(Ewv3to9_dementia_event)
390   ta E_age_group Ewv3to9_dementia_event, chi2 column row
391
392
393
394   /*
395   ---- SURVIVAL ANALYSIS IN COMPLETE DATA ----
396
397   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 E_time_event_dementia Ewv3to9_dementia_event ///
Ewv2_depressive_symptoms Ewv2_crp Ewv2_hdl_cholesterol ///
Ewv2_obesity_waist Ewv2_systolic_bp Ewv2_diastolic_bp ///
Ewv2_diabetes_diagnosed Ewv2_HbA1c ///
Ewv2_cardio2 Ewv2_cardio3 Ewv2_cardio4 Ewv2_cardio_number ///
Ewv2_waist_depress_group Ewv2_diabet_depress_group ///
Ewv2_hba1c_depress_group Ewv2_hdl_depress_group ///
Ewv2_sbp_depress_group Ewv2_dbp_depress_group ///
Ewv2_crp_depress_group Ewv2_ca2_depress_group ///
Ewv2_ca3_depress_group Ewv2_ca4_depress_group ///
E_age E_sex E_eduaction E_maritalstatus_4cat E_wealthquintiles ///
E_smoking_3cat E_alcohol_status E_cvd_comorbidity,compact



* Declare Data to be Survival Data
* Time to event: E_time_event_dementia (months)
* Censoring: Ewv3to9_dementia_event (1=dementia, 0=censored)
* Command is stset TIMETOEVENT, failure(CENSORVARIABLE)


stset E_time_event_dementia, failure (Ewv3to9_dementia_event==1) id(idauniq)


*describe survival data using commnad stsum

stsum

stsum, by(Ewv2_ca2_depress_group)



* Kaplan Meier Curve estimation

sts list

sts list, by(Ewv2_ca2_depress_group)



* Kaplan Meier Curve Plot

* no frills plot

sts graph

* with frills

sts graph, xtitle("Time in Months") ytitle("Survival Prob") ///
title("Kaplan Meier Curve")

```

```
466
467   * With Greenwood CI limits
468
469   sts graph, gwood legend(off) xtitle("Time in Months") ytitle("Survival Prob") ///
470   title("Kaplan Meier Curve")
471
472
473   * Group Kaplan-Meier Curve Estimation
474   * Command is sts graph, by(GROUPVAR) OPTION OPTION OPTION Note: Must have sorted by GROUPVAR first
475
476   sort Ewv2_ca2_depress_group
477
478   sts list, by(Ewv2_ca2_depress_group)
479
480   * graph with frills
481
482   sts graph, by(Ewv2_ca2_depress_group) xlabel(0(20)180) ylabel(0.80(.05)1) xtitle("Time in Months")
       ///
483   ytitle("Survival Prob") title("Kaplan Meier Curve")
484
485
486
487   * calculate person-time and incidence rates using command ststime
488
489   stptime,title(Person-years)
490
491   stptime, title(Person-years) per(1000)
492
493
494
495   /*
496
497   Repeat to find incident case per category
498
499   Ewv2_depressive_symptoms
500   Ewv2_crp
501   Ewv2_hdl_cholesterol
502   Ewv2_obesity_waist
503   Ewv2_systolic_bp
504   Ewv2_diastolic_bp
505   Ewv2_diabetes_diagnosed
506   Ewv2_HbA1c
507   Ewv2_cardio_number
508   Ewv2_cardio2
509
510
511   Ewv2_crp_depress_group
512   Ewv2_hdl_depress_group
513   Ewv2_waist_depress_group
514   Ewv2_sbp_depress_group
515   Ewv2_dbp_depress_group
516   Ewv2_diabet_depress_group
517   Ewv2_hba1c_depress_group
518   Ewv2_ca2_depress_group
519
520
521   */
522
523
524   ta Ewv2_ca2_depress_group
525
526   * calculate person-time  by category
527
528   stptime, by(Ewv2_ca2_depress_group)
529
530   stptime, by(Ewv2_ca2_depress_group) per(1000)
531
532
```

```
533
534   * mean and median of follow-up
535   sum E_time_event_dementia
536   sum E_time_event_dementia, detail
537
538
539
540
541
542
543   /* Log Rank Test of equality of survival distributions
544    (NULL: equality of survival distributions among groups)
545    We will consider including the predictor if the test has a p-value of 0.2 – 0.25 or less.
546    If the predictor has a p-value greater than 0.25 in a univariate analysis
547    it is highly unlikely that it will contribute anything to a model which includes other
      predictors.
548    Command is sts test GROUPVAR
549   */
550
551
552   sts test Ewv2_ca2_depress_group, logrank
553
554   sts test E_age, logrank
555
556   sts test E_sex, logrank
557
558   sts test E_eduaction, logrank
559
560   sts test E_maritalstatus_4cat, logrank
561
562   sts test E_wealthquintiles, logrank
563
564   sts test E_smoking_3cat, logrank
565
566   sts test E_alcohol_status, logrank
567
568   sts test E_cvd_comorbidity, logrank
569
570
571
572
573
574
575   /* Cox PH regression model
576
577   using the command stcox
578
579   --- Building the model ---
580
581   Model 1: unadjusted - single predictor of group
582   Model 2: model 1 + sociodemographics: age sex education marital status and wealth
583   Model 3: model 2 + lifestyle/health indicators: smoking, alcohol consumption, cvd comorbidity
584
585
586   !! I didn't adjust for physical activity because this variable can't be used in CHARLS (missing
      values)
587
588   */
589
590
591   * Unadjusted model - model 1 - single predictor
592
593   stcox Ewv2_ca2_depress_group
594
595   * define design var by using i.(by group)
596
597   stcox i.Ewv2_ca2_depress_group
598
```

```
599
600   * Adjusted models - multivariable Cox model
601   * controlling for covariates
602
603   * model 2: model 1 + adjust for sociodemographics: age sex education marital status and wealth
604
605   stcox i.Ewv2_ca2_depress_group E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.
      E_wealthquintiles
606
607
608   * model 3: model 2 + adjust for lifestyle/ health indicators
609
610   stcox i.Ewv2_depressive_symptoms E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.
      E_wealthquintiles ///
611   i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity
612
613
614
615
616
617   * Coefficients instead of hazard ratios by specifing the option nohr
618
619   stcox i.Ewv2_ca2_depress_group, nohr
620
621
622   stcox i.Ewv2_ca2_depress_group E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.
      E_wealthquintiles ///
623   i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity, nohr
624
625
626
627
628
629   * Multivariable model development
630   * Likelihood-ratio tests
631
632
633
634   *install eststo
635   findit eststo
636
637
638   * ---- rx controlling for age and sex -----*
639   quietly: stcox E_age i.E_sex
640   eststo modelagesex
641
642   quietly: stcox E_age i.E_sex i.Ewv2_ca2_depress_group
643   eststo modelagesex_4group
644
645   lrtest modelagesex modelagesex_4group
646
647
648
649   * ---- rx controlling for sociodemographics -----*
650   quietly: stcox E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles
651   eststo modelsociodemo
652
653   quietly: stcox E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles i.
      Ewv2_ca2_depress_group
654   eststo modelsociodemo_4group
655
656   lrtest modelsociodemo modelsociodemo_4group
657
658
659   * ---- rx controlling for lifestyle/health indicators-----*
660   quietly: stcox i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity
661   eststo modelcardiovascular
662
```

```
quietly: stcox i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity i.Ewv2_ca2_depress_group
eststo modelcardiovascular_4group

lrtest modelcardiovascular modelcardiovascular_4group



* side-by-side comparison of models


quietly: stcox i.Ewv2_ca2_depress_group
eststo model1


quietly: stcox E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles i.
Ewv2_ca2_depress_group
eststo model2


quietly: stcox E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity i.Ewv2_ca2_depress_group
eststo model3









* Display Betas and Summary Statistics
estout model1 model2 model3, stats(n chi2 bic, star(chi2)) prehead("Betas")

/* Key Interpretattion
Chi2 = Value of LR test comparing the model fit (“full”) to intercept only (“reduced”)
bic = Schwarz’ Bayesian Information Criterion = It is a function of the log-likelihood.
Smaller values indicate a better fit.
*/

* Display Hazard Ratios and Model Fit Statistics. Option eform produces hazard ratios
estout model1 model2 model3, eform stats(n chi2 bic, star(chi2)) prehead("Hazard Ratios")




* Postestimation tools for stcox

* Test of proportional hazards

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.Ewv2_ca2_depress_group E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.
E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity, nohr ///
tvc(Ewv2_ca2_depress_group E_age E_sex E_eduaction E_maritalstatus_4cat E_wealthquintiles ///
```

```
729   E_smoking_3cat E_alcohol_status E_cvd_comorbidity texp(ln(E_time_event_dementia))
730
731
732
733   /* Proportionality Assumption - method 2
734   by using the Schoenfeld and scaled Schoenfeld residuals
735   In the stphtest command we test the proportionality of the model as a whole
736   and by using the detail option we get a test of proportionality for each predictor.
737   By using the plot option we can also obtain a graph of the scaled Schoenfeld assumption.
738   If the tests in the table are not significance (p-values over 0.05)
739   then we can not reject proportionality and we assume
740   that we do not have a violation of the proportional assumption.
741   The stphplot command uses log-log plots to test proportionality
742   and if the lines  in these plots are parallel then we have further indication
743   that the predictors do not violate the proportionality assumption.
744   */
745
746   quietly stcox Ewv2_ca2_depress_group E_age E_sex E_eduaction E_maritalstatus_4cat
      E_wealthquintiles ///
747   E_smoking_3cat E_alcohol_status E_cvd_comorbidity, schoenfeld(sch*) scaledsch(sca*)
748   stphtest, detail
749   stphtest, plot(Ewv2_ca2_depress_group) msym(oh)
750   stphtest, plot(E_age) msym(oh)
751   stphtest, plot(E_sex) msym(oh)
752   stphtest, plot(E_eduaction) msym(oh)
753   stphtest, plot(E_maritalstatus_4cat) msym(oh)
754   stphtest, plot(E_wealthquintiles) msym(oh)
755   stphtest, plot(E_smoking_3cat) msym(oh)
756   stphtest, plot(E_alcohol_status) msym(oh)
757   stphtest, plot(E_cvd_comorbidity) msym(oh)
758
759
760
761
762
763   stphplot, by(Ewv2_ca2_depress_group) plot1(msym(oh)) plot2(msym(th))
764   stphplot, by(E_age) plot1(msym(oh)) plot2(msym(th))
765   stphplot, by(E_sex) plot1(msym(oh)) plot2(msym(th))
766   stphplot, by(E_eduaction) plot1(msym(oh)) plot2(msym(th))
767   stphplot, by(E_maritalstatus_4cat) plot1(msym(oh)) plot2(msym(th))
768   stphplot, by(E_wealthquintiles) plot1(msym(oh)) plot2(msym(th))
769   stphplot, by(E_smoking_3cat) plot1(msym(oh)) plot2(msym(th))
770   stphplot, by(E_alcohol_status) plot1(msym(oh)) plot2(msym(th))
771   stphplot, by(E_cvd_comorbidity) plot1(msym(oh)) plot2(msym(th))
772
773
774
775   * Assessment of PH Assumption: adjust for age and sex
776   stphplot, by(Ewv2_ca2_depress_group) adjust(E_age E_sex) nolntime plot1opts(symbol(none) color(
      black) lpattern(dash)) ///
777   plot2opts(symbol(none) color(green)) plot3opts(symbol(none) color(red)) ///
778   title("Assessment of PH Assumption") subtitle(" Predictor is Ewv2_ca4_depress_group") xtitle(
      "months")
779
780
781
782   * Assessment of PH Assumption: adjust for model 2
783   stphplot, by(Ewv2_ca2_depress_group) adjust(E_age E_sex E_eduaction E_maritalstatus_4cat
      E_wealthquintiles) ///
784   nolntime plot1opts(symbol(none) color(black) lpattern(dash)) ///
785   plot2opts(symbol(none) color(green)) plot3opts(symbol(none) color(red)) ///
786   title("Assessment of PH Assumption") subtitle(" Predictor is Ewv2_ca4_depress_group") xtitle(
      "months")
787
788
789
790   * Assessment of PH Assumption: adjust for model 3
791   stphplot, by(Ewv2_ca2_depress_group) adjust(E_age E_sex E_eduaction E_maritalstatus_4cat
```

```
    E_wealthquintiles ///
792 E_smoking_3cat E_alcohol_status E_cvd_comorbidity) ///
793 nolntime plot1opts(symbol(none) color(black) lpattern(dash)) ///
794 plot2opts(symbol(none) color(green)) plot3opts(symbol(none) color(red)) ///
795 title("Assessment of PH Assumption") subtitle(" Predictor is Ewv2_ca4_depress_group") xtitle(
    "months")
796
797
798
799
800
801 /* Test of overall goodness of fit
802 Goodness of fit of the final model
803 2 methods:
804  - by using the commnad stcoxgof (good fit = non sig p-value)
805  - by using the Cox-Snell residuals
806     to create the Nelson-Aalen cumulative hazard function
807     If the hazard function follows the 45 degree line then we know that it approximately
808     has an exponential distribution with a hazard rate of one and that the model fits the data
    well.
809     If the model fits the data, the plot of the cumulative hazard versus cs
810     should approximate a straight line with slope 1.
811 */
812
813
814 * by using the commnad stcoxgof
815
816 * install stcoxgof
817
818 findit stcoxgof
819
820
821 stcox Ewv2_ca2_depress_group E_age E_sex E_eduaction E_maritalstatus_4cat E_wealthquintiles ///
822 E_smoking_3cat E_alcohol_status E_cvd_comorbidity, mgale(mgale)
823
824
825 stcoxgof
826
827
828
829
830
831 * by using the Cox-Snell residuals
832
833 quietly stcox Ewv2_ca2_depress_group E_age E_sex E_eduaction E_maritalstatus_4cat
    E_wealthquintiles ///
834 E_smoking_3cat E_alcohol_status E_cvd_comorbidity
835 predict cs, csnell
836
837 * or
838
839 quietly stcox Ewv2_ca2_depress_group
840 predict cs, csnell
841
842
843 stset cs, failure(Ewv3to9_dementia_event)
844 sts generate km = s
845 generate H = -ln(km)
846 line H cs cs, sort ytitle("") clstyle(. refline)
847
848
849
850
851
852
853
854 /* Cox PH regression model for independent depressive symptoms and CA exposure variables
855
```

```
Ewv2_depressive_symptoms
Ewv2_crp
Ewv2_hdl_cholesterol
Ewv2_obesity_waist
Ewv2_systolic_bp
Ewv2_diastolic_bp
Ewv2_diabetes_diagnosed
Ewv2_HbA1c
Ewv2_cardio_number
Ewv2_cardio2
Ewv2_cardio3
Ewv2_cardio4

*/



stset E_time_event_dementia, failure (Ewv3to9_dementia_event==1) id(idauniq)


* Unadjusted model 1

stcox i.Ewv2_depressive_symptoms


* Adjusted models - multivariable Cox model
* controlling for covariates

* model 2: model 1 + adjust for sociodemographics: age sex education marital status and wealth

stcox i.Ewv2_depressive_symptoms E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.
E_wealthquintiles

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

stcox i.Ewv2_depressive_symptoms E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.
E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity


* repeat for each independent variable from the list above



** ---------------------- **



/* MULTIPLE IMPUTATION (MI)

To handle with missing values of covariates


useful sources for MI and MICE:

https://stats.idre.ucla.edu/stata/seminars/mi_in_stata_pt1_new/
https://www.stata.com/manuals/mi.pdf - see page 139
https://www.stata.com/meeting/switzerland16/slides/medeiros-switzerland16.pdf
https://www.youtube.com/watch?v=i6SOlq0mjuc&ab_channel=StataCorpLLC
https://dss.princeton.edu/training/MIStata.pdf



```

```
Preparing to conduct MI
1. examine the number and proportion of missing values among the variables of interest
    use the mdesc command
2. examine missing data patterns
    use commands mi set and mi misstable patterns
3. identify potential auxiliary variables


Run MI using chained equations (MICE)
using the commands
1. how (in what style) to store the imputations
 mi set wide
2. which variables will be imputed
 mi register imputed
3. optionally, which variables will not be imputed
 mi register regular
4. what imputation method is implemented to impute each of var - MICE
 mi impute chained

*/




/*

1. examining missing values
    install packages:
    * install mdesc
    * install tabmiss
    * insatll dm31
    * insall mvpatterna

*/

search mdesc
search rmiss2
search mvpatterns


* examining number of missing values vs non-missing in each variable

mdesc E_age E_sex E_eduaction E_maritalstatus_4cat E_wealthquintiles ///
E_smoking_3cat E_physicalactivity E_alcohol_status E_cvd_comorbidity E_memory_wordrecall




* examining missing data patterns

mi set wide

mi misstable summarize E_age E_sex E_eduaction E_maritalstatus_4cat ///
E_wealthquintiles E_smoking_3cat E_physicalactivity ///
E_alcohol_status E_cvd_comorbidity ///

mi misstable patterns E_age E_sex E_eduaction ///
E_maritalstatus_4cat E_wealthquintiles ///
E_smoking_3cat E_physicalactivity E_alcohol_status E_cvd_comorbidity


/*
 identifying potential auxiliary var
* Auxiliary variables are either correlated with a missing variable(s)
(the recommendation is r > 0.4) or are believed to be associated with missingness
- a priori knowledge of var that would make good auxiliary var
- identify potential candidates by examining associations between missing var and other var in
```

```
    the dataset
        running correlation using the command: pwcorr v1 v2 v3, obs
        the recommnedation for good correlation is r > 0.4


    Missing var to be imputed:

        E_eduaction E_wealthquintiles
        E_smoking_3cat E_physicalactivity E_alcohol_status




    Potential auxiliary var:
    DV:  Ewv3to9_dementia_event
    IV:  Ewv2_depressive_symptoms Ewv2_crp Ewv2_hdl_cholesterol
    Ewv2_obesity_waist Ewv2_systolic_bp Ewv2_diastolic_bp
    Ewv2_diabetes_diagnosed Ewv2_HbA1c
    other var:
        E_age E_sex E_maritalstatus_4cat E_cvd_comorbidity

    */


    * correlation

    pwcorr E_eduaction E_wealthquintiles ///
        E_smoking_3cat E_physicalactivity E_alcohol_status ///
        Ewv3to9_dementia_event Ewv2_depressive_symptoms Ewv2_crp Ewv2_hdl_cholesterol ///
        Ewv2_obesity_waist Ewv2_systolic_bp Ewv2_diastolic_bp ///
        Ewv2_diabetes_diagnosed Ewv2_HbA1c ///
        E_age E_sex E_maritalstatus_4cat E_cvd_comorbidity, obs


    /* The correlation showed that all the following var are good auxiliary:
        Ewv3to9_dementia_event Ewv2_depressive_symptoms Ewv2_crp Ewv2_hdl_cholesterol
        Ewv2_systolic_bp Ewv2_diabetes_diagnosed E_age E_sex E_cvd_comorbidity
    A good auxiliary does not have to be correlated with every variable to be useful
    And it's not problematic if it has missing info of it's own
    */



    /*
    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.
```

```
1057
1058
1059    IMPUTATION PHASES
1060
1061    1. mi set wide
1062        style to store imputations
1063
1064    2. mi register imputed
1065        identifies which variables in the imputation model have missing information.
1066
1067    3. mi register regular (! optional)
1068        which variables will not be imputed
1069
1070    4. mi impute chained
1071        where the user specifies the imputation model to be used
1072        and the number of imputed datasets to be created.
1073        Example:
1074            mi impute chained (regress) bmi age (logit) female ///
1075            (mlogit) race = bpdiast i.region, add(20)
1076
1077    5. mi estimate
1078        is used as a prefix to the standard regress command.
1079        This executes the specified estimation model within each of the 20 imputed datasets
1080        to obtain 20 sets of coefficients and standard errors.
1081        Stata then combines these estimates to obtain one set of inferential statistics.
1082        In the output from mi estimate you will see some metrics: Imputation Diagnostics
1083        information for RVI (Relative Increase in Variance),
1084        FMI (Fraction of Missing Information),
1085        DF (Degrees of Freedom) ,
1086        RE (Relative Efficiency),
1087        and the between imputation and the within imputation variance estimates
1088        to examine how the standard errors (SEs) are calculated.
1089
1090
1091    ----------------
1092
1093    SELECTING MY IMPUTATION MODEL
1094
1095    - MICE -> mi impute chained
1096
1097    - var to be imputed:
1098
1099        linear regression for continuous var (regress) -> none
1100
1101        logistic for the binary var (logit) -> none
1102
1103        multinomial logistic for our nominal categorical var (mlogit) ->
1104        E_eduaction E_wealthquintiles
1105        E_smoking_3cat E_physicalactivity E_alcohol_status
1106
1107
1108
1109    - auxiliary var:
1110
1111        DV -> Ewv3to9_dementia_event
1112        IV -> Ewv2_depressive_symptoms Ewv2_crp Ewv2_hdl_cholesterol
1113        Ewv2_systolic_bp Ewv2_diabetes_diagnosed
1114        other covariates -> E_age E_sex E_cvd_comorbidity
1115
1116
1117
1118    - imputation numbers (m) -> 10
1119
1120        White et al. (2010) recommendation: use the rule that m should equal the percentage of
        incomplete cases
1121
1122
1123    - rseed (53421) for reproducability reasons
```

```
1124
1125
1126    - (! OPTIONAL) advance impute options -> force
1127
1128        proceed with imputation, even when missing imputed values (e.g. auxiliary have missing data)
        are encountered
1129
1130    - impute options -> savetrace (trace1)
1131
1132        specifies Stata to save the means and standard deviations of imputed values from each
        iteration to a Stata dataset named “trace1
1133    */
1134
1135
1136    mi set wide
1137
1138
1139    mi register imputed E_eduaction E_wealthquintiles ///
1140        E_smoking_3cat E_physicalactivity E_alcohol_status
1141
1142
1143
1144    mi impute chained (mlogit) E_eduaction E_wealthquintiles E_smoking_3cat E_physicalactivity
        E_alcohol_status = Ewv3to9_dementia_event Ewv2_depressive_symptoms Ewv2_crp Ewv2_hdl_cholesterol
        Ewv2_systolic_bp Ewv2_diabetes_diagnosed E_age E_sex E_cvd_comorbidity, add(10) rseed(53421)
        savetrace(trace1)
1145
1146
1147
1148
1149    * save imputed data
1150
1151    * plot imputations
1152
1153
1154    *it will open a file named trace1
1155    use trace1, clear
1156
1157    describe
1158
1159
1160    reshape wide *mean *sd, i(iter) j(m)
1161
1162    tsset iter
1163
1164
1165
1166
1167    /*
1168    The trace plot below graphs the predicted means value produced during the first imputation chain.
1169    As before, the expectations is that the values would vary randomly to incorporate variation into
        the predicted values for read.
1170    */
1171
1172    tsline E_eduaction_mean1, name(mice1,replace)legend(off) ytitle("Mean of education")
1173    tsline E_wealthquintiles_mean1, name(mice1,replace)legend(off) ytitle("Mean of wealth")
1174    tsline E_smoking_3cat_mean1, name(mice1,replace)legend(off) ytitle("Mean of smoking")
1175    tsline E_physicalactivity_mean1, name(mice1,replace)legend(off) ytitle("Mean of physical activity")
1176    tsline E_alcohol_status_mean1, name(mice1,replace)legend(off) ytitle("Mean of alcohol status")
1177
1178
1179    /*
1180
1181    All 10 imputation chains can also be graphed simultaneously to make sure that nothing unexpected
        occurred in a single chain.
1182    Every chain is obtained using a different set of initial values and this should be unique.
1183    Each colored line represents a different imputation.
1184    So all 10 imputation chains are overlaid on top of one another.
```

```
*/

tsline E_alcohol_status_mean*, name(mice1,replace)legend(off) ytitle("Mean of alcohol")
tsline E_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 E_time_event_dementia, failure (Ewv3to9_dementia_event==1) id(idauniq)

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




/*
Independent risk factors

Ewv2_depressive_symptoms
Ewv2_crp
Ewv2_hdl_cholesterol
Ewv2_obesity_waist
Ewv2_systolic_bp
Ewv2_diastolic_bp
Ewv2_diabetes_diagnosed
Ewv2_HbA1c
Ewv2_cardio_number
Ewv2_cardio2

*/



* Depressive symptoms

* Unadjusted model - model 1 - single predictor

* Model 1 (default coefficents)
mi estimate: stcox Ewv2_depressive_symptoms

* Model 1: define design var by using i.
mi estimate: stcox i.Ewv2_depressive_symptoms


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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_depressive_symptoms



* Adjusted models - multivariable Cox model
* controlling for covariates

```

```
* model 2: model 1 + adjust for sociodemographics: age sex education marital status and wealth

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_depressive_symptoms ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles


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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_depressive_symptoms ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity



* repeat for each independent variable from the list above




/*
Combined effects Cox regression models

Ewv2_crp_depress_group
Ewv2_hdl_depress_group
Ewv2_waist_depress_group
Ewv2_sbp_depress_group
Ewv2_dbp_depress_group
Ewv2_diabet_depress_group
Ewv2_hba1c_depress_group
Ewv2_ca2_depress_group

*/



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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group


* Adjusted models - multivariable Cox model
* controlling for covariates

* model 2: model 1 + adjust for sociodemographics: age sex education marital status and wealth

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles

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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity


* repeat for each variable from the list above



*-------------------------*



/*

```

```
*** SENSITIVITY ANALYSES ***


1) interarction effect of gender and age_group

2) survival analysis stratified by age
two age groups: young old <70 and old old >=70

3) depressive symptoms as continuous variable
and >= 3 and >=4 cardiometabolic multimorbidity

4) exclude participants with cvd

5) survival analysis without IQCODE

6) Complete data

7) survival analysis limiting to 5 year follow-up period


Repeat on all independent and combined variables

Ewv2_depressive_symptoms
Ewv2_crp
Ewv2_hdl_cholesterol
Ewv2_obesity_waist
Ewv2_systolic_bp
Ewv2_diastolic_bp
Ewv2_diabetes_diagnosed
Ewv2_HbA1c
Ewv2_cardio_number
Ewv2_cardio2


Ewv2_crp_depress_group
Ewv2_hdl_depress_group
Ewv2_waist_depress_group
Ewv2_sbp_depress_group
Ewv2_dbp_depress_group
Ewv2_diabet_depress_group
Ewv2_hba1c_depress_group
Ewv2_ca2_depress_group


*/




* 1) Interaction effect

* sex*risk factor

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_cardio2 i.E_sex#i.Ewv2_cardio2


mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_cardio2 ///
E_age i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity ///
i.E_sex#i.Ewv2_cardio2


* age*risk factor

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_cardio2 c.E_age#i.Ewv2_cardio2


```

```
mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_cardio2 ///
E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity ///
 c.E_age#i.Ewv2_cardio2



/* 2) Survival analysis stratified by age

generate age group variable
Age groups: 1) young old (< 70) 2) old old (>= 70)

Kaplan Meier curves
Cox regression models in imputed data

young old <70
if E_age_group==1

old old >70
if E_age_group==2

*/



gen E_age_group=1 if E_age < 70
replace E_age_group=2 if E_age >=70 & !missing(E_age)

label var E_age_group "Age groups <70 young-old / 70 old-old"
lab def age_group 1 "young old <70" 2 "old old >70"
lab val E_age_group age_group

tab E_age_group



* COX PH REGRESSION MODEL IN IMPUTED DATASET


* Declare Data to be Survival Data by using mi

mi stset E_time_event_dementia, failure (Ewv3to9_dementia_event==1) id(idauniq)



* YOUNG OLD <70 Cox regression models

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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group if E_age_group==1


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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group ///
i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity if E_age_group==1



* OLD OLD >70 Cox regression models


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

```

```
mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group if E_age_group==2


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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group ///
i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity if E_age_group==2









* 3) On depressive symptom continuous variable


* COX PH REGRESSION MODEL IN IMPUTED DATASET

* Declare Data to be Survival Data by using mi

mi stset E_time_event_dementia, failure (Ewv3to9_dementia_event==1) id(idauniq)



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

mi estimate, eform("Haz. Ratio"): stcox Ewv2_cesd_score


* 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 Ewv2_cesd_score ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles


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

mi estimate, eform("Haz. Ratio"): stcox Ewv2_cesd_score ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity









* Cardiometabolic multimorbidty >= 3


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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_cardio3


* 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.Ewv2_cardio3 ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles


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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_cardio3 ///
```

```
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity




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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_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.Ewv2_ca3_depress_group ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles


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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca3_depress_group ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_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.Ewv2_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.Ewv2_cardio4 ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles


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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_cardio4 ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity




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

mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_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.Ewv2_ca4_depress_group ///
E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles


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

```

```
1593  mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca4_depress_group ///
1594  E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
1595  i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity
1596
1597
1598
1599
1600
1601
1602  /*
1603
1604  4) exclude participants with cvd
1605
1606  use the command if E_cvd_comorbidity==0
1607
1608  */
1609
1610
1611  * COX PH REGRESSION MODEL IN IMPUTED DATASET
1612
1613
1614  * Declare Data to be Survival Data by using mi
1615
1616  mi stset E_time_event_dementia, failure (Ewv3to9_dementia_event==1) id(idauniq)
1617
1618
1619  * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1620
1621  mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group if E_cvd_comorbidity==0
1622
1623
1624
1625  * Model 3: model 2 + adjust for lifestyle/health indicators
1626
1627  mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group ///
1628  E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
1629  i.E_smoking_3cat i.E_alcohol_status if E_cvd_comorbidity==0
1630
1631
1632
1633
1634
1635  * 5) Exclude dementia cases identified with IQCODE
1636
1637
1638  * Declare Data to be Survival Data by using mi
1639
1640  mi stset E_time_event_dementia_report_no_, failure (Ewv3to9_dementia_event_no_iqcode==1) id(
      idauniq)
1641
1642
1643
1644  * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1645
1646
1647  mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group
1648
1649
1650  * Model 3: model 2 + adjust for cvd health
1651
1652  mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group ///
1653  E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
1654  i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity
1655
1656
1657
1658  * 6) On complete cases (see above)
1659
```

```
1660
1661
1662    /*
1663    7) survival analysis limiting to 5 year follow-up period
1664
1665    elsa follow-up wave 3-6
1666
1667    */
1668
1669
1670
1671    merge 1:m RAHHIDPN using "S:\Research\pkstudies\Study3_cardio_lca\HRS\hrs_lca data sensitivity
        9to12followup.dta"
1672
1673
1674
1675    * Declare Data to be Survival Data by using mi
1676
1677    mi stset Ewv3to6_time_event_dementia, failure (Ewv3to6_dementia_event==1) id(idauniq)
1678
1679
1680    * Model 1 ask for hazard ratio by using the option eform("Haz.Ratio")
1681
1682    mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group
1683
1684
1685    * Adjusted models - multivariable Cox model
1686    * controlling for covariates
1687
1688    * Model 2: model 1 + adjust for socio-demographics: age sex education marital status and wealth
1689    mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group ///
1690    E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles
1691
1692
1693
1694
1695    * Model 3: model 2 + adjust for lifestyle/health indicators
1696
1697    mi estimate, eform("Haz. Ratio"): stcox i.Ewv2_ca2_depress_group ///
1698    E_age i.E_sex i.E_eduaction i.E_maritalstatus_4cat i.E_wealthquintiles ///
1699    i.E_smoking_3cat i.E_alcohol_status i.E_cvd_comorbidity
1700
1701
1702
1703
1704
1705
1706
1707    * ------------------------------ *
1708
1709
1710
1711
1712
1713
1714
```