

```

1
2  version 15.1
3
4  #####
5  #####
6  **CLEANING**
7
8  cls
9  clear
10 set more off
11
12 /*initial cleaning*/
13 use "db0", clear
14 distinct n_eid
15 describe, short
16 drop n_20001_*_* n_20002_1_* n_20002_2_* n_20003_*_* n_20006_*_* n_20007_*_* n_20008_*_* /*
17 */ n_20009_*_* n_20009_*_* n_20015_*_* n_20043_*_* n_20044_*_* n_20058_*_* n_20059_*_* /*
18 */ n_20060_*_* n_21_*_* n_22* n_2644* n_2654* n_28* n_30* n_34????? n_365* n_368* n_37* n_400* n_41* /*
19 */ n_44* n_53* n_595* n_5985* n_5986* n_5992* n_5993* n_6014_0_0-n_6017_1_0 n_6024* /*
20 */ n_6143* n_6144* n_6158* n_6162* n_6164_1* n_6164_2* n_6183* n_6194* n_757* n_777* n_796* n_806* /*
21 */ n_816* n_826* n_845* n_94* n_90013_0_0-n_90050_0_0 n_90052_0_0-n_90126_0_0 /*
22 */ n_90128_0_0-n_110005_0_0 n_9006* s_2* s_40002* s_40006* s_40013* s_5987* s_5988* /*
23 */ ts_900* ts_40005_* n_87_* n_2624_1_0 n_2624_2_0 n_2634_1_0 n_2634_2_0
24 describe, short
25
26 drop n_*_1_* n_*_2_* n_10855* n_1090* n_110* n_1279* n_1478_0_0 n_1488_0_0 n_1498_0_0 /*
27 */ n_1508_0_0 n_1518_0_0 n_1528_0_0 n_1538_0_0 n_1548_0_0 n_1568_0_0 n_1578_0_0 n_1588_0_0 /*
28 */ n_1598_0_0 n_1608_0_0 n_1618_0_0 n_21022_0_0 n_51_0_0 n_6138_0_1 n_6138_0_2 n_6138_0_3 /*
29 */ n_6138_0_4 n_6138_0_5 n_6142_0_1 n_6142_0_2 n_6142_0_3 n_6142_0_4 n_6142_0_5 n_6142_0_6 /*
30 */ n_767_0_0 n_84_0_0 n_84_0_1 n_84_0_2 n_84_0_3 n_84_0_4 n_84_0_5 s_6023_1_0 s_6025_1_0
31 describe, short
32
33 foreach i of numlist 1009674 1105541 1161338 1165508 1195587 1446958 1478377 1545979 1667988 1669836 1751658 1860814 1920695
1986543 2068378 2070413 2190001 2278762/*
34 */ 2351481 2380966 2625508 2629413 2662650 2667559 2775056 2788405 3007998 3125030 3229075 3422982 3618705
3768608 3868920 3904085 4100907 4121914/*
35 */ 4387267 4526671 4724492 4785314 4980534 5145670 5163347 5197631 5211598 5399765 5557036 5596258 5617445
5634380 5642632 5812050 5862068 5895219/*
36 */ 5976888 6016642 {
37 drop if n_eid == `i'
38 } /*as per Biobank indication email, 26/10/2018, patients withdrawn*/
39 describe, short
40
41 preserve
42 describe, replace clear
43 export excel name varlab using "labs.xlsx", firstrow(variables) replace
44 display "$S_TIME $S_DATE"
45 restore
46

```

```

47 ukbrename, dictionary("labs1.csv") nostdcheck
48 drop currempl
49 drop date0_rep*
50 mdesc ddate*
51 egen float miss_dd = rowmiss(ddate*)
52 tab miss_dd
53 tabstat ddate*, statistics(count) by(miss_dd)
54 count if (ddate == ddate_rep1 & miss_d == 1)
55 list n_eid ddate ddate_rep1 if (ddate != ddate_rep1 & miss_d == 1)
56 replace ddate = ddate_rep1 if (ddate != ddate_rep1 & miss_d == 1)
57 list n_eid ddate ddate_rep1 if n_eid == 1276410
58 count if (ddate == ddate_rep1 == ddate_rep2) & miss_d == 0
59 egen byte diff = diff(ddate*) if miss_d == 0
60 tab diff
61 drop miss_dd ddate_r* diff
62 drop death_1stICD10_r*
63 rename n_eid rep1 n_eid
64 describe, short
65 save "db1", replace
66 display "$S_TIME $S_DATE"
67
68 *rectangular random numbers for day of birth, according to months and leap years
69 cls
70 use "db1", clear
71 set seed 151118
72 generate birth_day = floor((30)*runiform() + 1) if (bmonth == 4 | bmonth == 6 | bmonth == 9 | bmonth == 11)
73 replace birth_day = floor((31)*runiform() + 1) if (bmonth == 1 | bmonth == 3 | bmonth == 5 | bmonth == 7 | bmonth == 8 | bmonth ==
    10 | bmonth == 12)
74
75 sum byear
76 gen leapyear = 1 if (byear == 1936 | byear == 1940 | byear == 1944 | byear == 1948 | byear == 1952 | byear == 1956 | byear
    == 1960 | byear == 1964 | byear == 1968)
77 replace birth_day = floor((29)*runiform() + 1) if (bmonth == 2 & leapyear == 1)
78 replace birth_day = floor((28)*runiform() + 1) if (bmonth == 2 & leapyear == .)
79 misstable summarize birth_day
80 drop leapyear
81 histogram birth_day
82 graph close _all
83
84 *dates in numbers
85 tostring byear, replace
86 tostring bmonth, replace force
87 tostring birth_day, replace
88 gen str birth_date = birth_day + "/" + bmonth + "/" + byear
89 gen birth_d = date(birth_date, "DMY")
90 gen visit_d = date0
91 gen death_d = ddate
92
93 *centre label
94 tostring centre, replace
95 tab centre

```

```

96  replace centre = "Stockport"      if centre == "10003"
97  replace centre = "Manchester"     if centre == "11001"
98  replace centre = "Oxford"         if centre == "11002"
99  replace centre = "Cardiff"        if centre == "11003"
100 replace centre = "Glasgow"        if centre == "11004"
101 replace centre = "Edinburgh"     if centre == "11005"
102 replace centre = "Stoke"          if centre == "11006"
103 replace centre = "Reading"        if centre == "11007"
104 replace centre = "Bury"           if centre == "11008"
105 replace centre = "Newcastle"      if centre == "11009"
106 replace centre = "Leeds"          if centre == "11010"
107 replace centre = "Bristol"        if centre == "11011"
108 replace centre = "Barts"          if centre == "11012"
109 replace centre = "Nottingham"     if centre == "11013"
110 replace centre = "Sheffield"      if centre == "11014"
111 replace centre = "Liverpool"      if centre == "11016"
112 replace centre = "Middlesbrough"  if centre == "11017"
113 replace centre = "Hounslow"       if centre == "11018"
114 replace centre = "Croydon"        if centre == "11020"
115 replace centre = "Birmingham"    if centre == "11021"
116 replace centre = "Swansea"        if centre == "11022"
117 replace centre = "Wrexham"        if centre == "11023"
118
119 *censoring: https://biobank.ctsu.ox.ac.uk/crystal/exinfo.cgi?src=Data\_providers\_and\_dates
120 tab centre, missing
121 gen country = "sco" if (centre == "Edinburgh" | centre == "Glasgow")
122 replace country = "wal" if (centre == "Cardiff" | centre == "Swansea")
123 replace country = "eng" if (country == "")
124 tab country, missing
125
126 cls
127 tab date0
128 cls
129 tab ddate
130 tab ddate if country != "sco"
131 local x = mdy(1,31,2016)
132 replace death_d = . if (death_d > `x' & country != "sco")
133 tabstat ddate death_d, statistics(count) by(country)
134 bys country: list ddate death_d if (ddate !=. & death_d == .)
135 tabstat n_eid if (ddate !=. & death_d == .), statistics(count) by(country)
136
137 tab ddate if country == "sco"
138 local x = mdy(11,30,2015)
139 replace death_d = . if (death_d > `x' & country == "sco")
140 tabstat ddate death_d, statistics(count) by(country)
141 bys country: list ddate death_d if (ddate !=. & death_d == .)
142 tabstat n_eid if (ddate !=. & death_d == .), statistics(count) by(country)
143
144 *time scale: entry to exit
145 describe, short
146 gen exit_d = mdy(1,31,2016) if country != "sco"

```

```

147 replace exit_d      = mdy(11,30,2015) if country == "sco"
148 gen all_death       = 1 if death_d !=.
149 replace all_death    = 0 if death_d ==.
150 tab all_death
151 gen time             = (exit_d - visit_d)/365.24 if all_death == 0
152 replace time         = (death_d - visit_d)/365.24 if all_death == 1
153 sum time
154 count if time == 0
155 drop if time <0
156 distplot time, by(all_death)
157 graph close all
158 describe, short
159
160 *cause of death
161 gsort- death_1stICD10
162 cls
163 tab death 1stICD10          if all_death == 1, missing sort
164 mdesc death_1stICD10       if all_death == 1
165 gen cause_death           = "cancer" if (all_death == 1 & strpos(death_1stICD10,"C"))
166 replace cause_death       = "cvd"    if (all_death == 1 & strpos(death_1stICD10,"I"))
167 replace cause_death       = "other"   if (all_death == 1 & strpos(death_1stICD10,"I8"))
168 replace cause_death       = "other"   if (all_death == 1 & cause_death == "")
169 tab cause_death
170 gen cancer_death          = 1         if (all_death == 1), missing sort
171 replace cancer_death      = 0         if (cause_death == "cancer")
172 tab cancer_death
173 gen cvd_death             = 1         if cause_death == "cvd"
174 replace cvd_death         = 0         if cvd_death == .
175 tab cvd_death
176
177 *covariates manipulation
178 tab sex, missing
179 tab ethn, missing sort
180 sdecode ethn, replace
181 replace ethn = "White" if (ethn == "White" | ethn == "British" | ethn == "Irish" | ethn == "Any other white background")
182 replace ethn = "" if (ethn == "Prefer not to answer" | ethn == "Do not know")
183 replace ethn = "Non-White" if (ethn != "White" & ethn != "")
184 tab ethn, missing
185 sencode ethn, gsort(-eth) replace
186 labelbook ethn
187 tab ndrugs, missing
188 tab smok, missing
189 sdecode smok, replace
190 replace smok = "" if smok == "Prefer not to answer"
191 tab smok, missing
192 replace smok = "Former" if smok == "Previous"
193 sencode smok, gsort(-smok) replace
194 labelbook smok
195 tab nca, missing
196 gen bcancer             = "no"       if (nca == 0)
197 replace bcancer         = "yes"      if (nca != 0 & nca !=.)

```

```

/* all deaths with ICD10*/
/* C* = cancer */
/* I00-I79 = cvd */
/* no I8* as cvd */

```

```

198 tab bcancer, missing
199 sencode bcancer, gsort(bcancer) replace
200 labelbook bcancer
201 tab nnca, missing
202
203 **cvd - http://biobank.ctsu.ox.ac.uk/crystal/coding.cgi?id=6
204 /*
205 1066 heart/cardiac problem
206 1067 peripheral vascular disease
207 1074 angina
208 1075 heart attack/myocardial infarction
209 1076 heart failure/pulmonary odema
210 1079 cardiomyopathy
211 1588 hypertrophic cardiomyopathy (hcm / hocm)
212 1591 aortic aneurysm rupture
213 1592 aortic dissection
214 1492 aortic aneurysm
215 1087 leg claudication/ intermittent claudication
216
217 1081 stroke
218 1082 transient ischaemic attack (tia)
219 1083 subdural haemorrhage/haematoma
220 1086 subarachnoid haemorrhage
221 1491 brain haemorrhage
222 1583 ischaemic stroke
223
224 1222 type 1 diabetes
225 1223 type 2 diabetes
226
227 1192 renal/kidney failure
228 1193 renal failure requiring dialysis
229 1194 renal failure not requiring dialysis
230 */
231
232 foreach var of varlist n_20002_0_0 - n_20002_0_28 {
233     gen cvd_`var' = 1 if inlist(`var', 1066, 1067, 1074, 1075, 1076, 1079, 1588, 1591, 1592, 1492, 1087)
234     gen cva_`var' = 1 if inlist(`var', 1081, 1082, 1083, 1086, 1491, 1583)
235     gen ckd_`var' = 1 if inlist(`var', 1192, 1193, 1194)
236     gen t1d_`var' = 1 if `var' == 1222
237     gen t2d_`var' = 1 if `var' == 1223
238 }
239 egen float bcvd = rowtotal(cvd_n*)
240 egen float bcva = rowtotal(cva_n*)
241 egen float bckd = rowtotal(ckd_n*)
242 egen float bt1d = rowtotal(t1d_n*)
243 egen float bt2d = rowtotal(t2d_n*)
244 drop cvd_n* cva_n* ckd_n* t1d_n* t2d_n*
245
246 cls
247 foreach var of varlist bcvd-bt2d {
248     tab `var', sort m

```

```
249     replace `var' = 1 if `var'>0
250     tab `var', sort m
251 }
252 drop n_20002*
253 replace age0 = (visit_d - birth_d)/365.24
254
255 histogram bmi
256 mdesc bmi
257 gen lnbmi = ln(bmi)
258 histogram lnbmi
259
260 tab alcf, m
261 tab alcf, m nolab
262 replace alcf = . if alcf == -3
263 tab alcf, m nolab
264 *healthy score Alcohol
265 gen score_alc = 0 if alcf <3
266 replace score_alc = 1 if alcf >=3 & alcf !=.
267 tab score_alc, m
268
269 tab tv, m
270 tab tv, m nolab
271 replace tv = . if (tv == -3 | tv == -1)
272 *healthy score TV
273 gen score_tv = 1 if tv <3 | tv == -10
274 replace score_tv = 0 if tv >=3 & tv !=.
275 tab score_tv, m
276
277 replace tv = 0.5 if tv == -10
278 histogram tv
279 sum tv
280 replace tv = tv + 1
281 gen lntv = ln(tv)
282 histogram lntv
283
284 tab cpu, m
285 tab cpu, m nolab
286 replace cpu = 0.5 if cpu == -10
287 replace cpu = . if (cpu == -3 | cpu == -1)
288 histogram cpu
289 sum cpu
290 *replace cpu = cpu + 1
291 *gen lncpu = ln(cpu)
292 *histogram lncpu
293
294 tab sleep, m
295 tab sleep, m nolab
296 replace sleep = . if (sleep == -3 | sleep == -1)
297 *healthy score sleep
298 gen score_sleep = 0 if (sleep<6 | sleep>8 & sleep !=.)
299 replace score_sleep = 1 if (score_sleep == . & sleep !=.)
```

```

300  tab score_sleep, m
301
302  histogram sleep
303  sum sleep
304  *gen lnsle = ln(sleep)
305  *histogram lnsle
306  graph close _all
307
308  *Physical activity
309  order plew dplew pleexerc dpleexerc plesport dplesport fdidy ddiy wp, after(dv10)
310  cls
311  foreach var of varlist w10 m10 v10 {                               /*see
http://biobank.ctsu.ox.ac.uk/crystal/docs/TouchscreenQuestionsMainFinal.pdf, page 10, for details*/
312      tab `var', sort m
313      replace `var' = . if (inlist(`var', -1, -3, .))
314      replace `var' = 0 if (inlist(`var', -2))
315      tab `var', sort m
316  }
317  foreach var of varlist dw10 dm10 dv10 {
318      tab `var', sort m
319      replace `var' = . if (inlist(`var', -1, -3, .))
320      tab `var', sort m
321  }
322  foreach var of varlist w10 m10 v10 {
323      replace d`var' = 0 if `var' == 0
324      replace `var' = 0 if d`var' == 0
325  }
326  mdesc w10-dv10
327  gen totw = w10*dw10
328  gen totm = m10*dm10
329  gen totv = v10*dv10
330  mdesc totw totm totv
331  order totw totm totv, after(dv10)
332
333  foreach var of varlist totw totm totv {
334      xtilew `var'3 = `var', within(sex) nq(3)
335  }
336
337  *Leisure-time PA
338  foreach var of varlist plew pleexerc plesport {
339      tab `var', m
340      tab `var', m nolab
341      replace `var' = . if (`var' == -1 | `var' == -3)
342      tab `var', m nolab
343      gen n`var' = 0.25 if `var' == 1 /*weekly, using mid-point if interval; see
http://biobank.ctsu.ox.ac.uk/crystal/docs/TouchscreenQuestionsMainFinal.pdf, page 12, for details*/
344      replace n`var' = 0.625 if `var' == 2
345      replace n`var' = 1 if `var' == 3
346      replace n`var' = 2.5 if `var' == 4
347      replace n`var' = 4.5 if `var' == 5
348      replace n`var' = 7 if `var' == 6

```

```

349     replace n`var' = .      if `var' == .
350     tab n`var'
351
352     tab d`var', m
353     tab d`var', m nolab
354     replace d`var' = .      if (d`var' == -1 | d`var' == -3)      /*minutes, using mid-point if interval; see
http://biobank.ctsu.ox.ac.uk/crystal/docs/TouchscreenQuestionsMainFinal.pdf, page 13, for details*/
355     replace d`var' = 7.5    if d`var' == 1
356     replace d`var' = 22.5  if d`var' == 2
357     replace d`var' = 45    if d`var' == 3
358     replace d`var' = 75    if d`var' == 4
359     replace d`var' = 105   if d`var' == 5
360     replace d`var' = 150   if d`var' == 6
361     replace d`var' = 210   if d`var' == 7
362     tab d`var', m nolab
363 }
364 drop plew pleexerc plesport
365
366 tab dplew nplew, m
367 gen      plew_t = nplew*dplew, after(dplew)
368 replace plew_t = 0 if (nplew ==. & dplew ==.) & (typepa_a0 != . & typepa_a0 != -3) /*no answer means zero for this category with
values for other categories if the answer to the initial question was not "prefer not to answer (-3)" or missing*/
369
370 tab dpleexerc npleexerc, m
371 gen      pleexerc_t = npleexerc*dpleexerc, after(dpleexerc)
372 replace pleexerc_t = 0 if (npleexerc ==. & dpleexerc ==.) & (typepa_a0 != . & typepa_a0 != -3)
373
374 mdesc plew_t pleexerc_t
375 gen miss_plew      = 1 if plew_t      == .
376 replace miss_plew  = 0 if plew_t      != .
377 gen miss_pleexerc  = 1 if pleexerc_t == .
378 replace miss_pleexerc = 0 if pleexerc_t != .
379 tab miss_plew miss_pleexerc
380
381 gen score_pat = plew_t + pleexerc_t
382 mdesc score_pat
383 replace score_pat = plew_t      if (plew_t != . & pleexerc_t == .)
384 replace score_pat = pleexerc_t if (plew_t == . & pleexerc_t != .)
385 mdesc score_pat
386 distplot score_pat
387 gen score_pa      = 1 if (score_pat >=150 & score_pat !=.)
388 replace score_pa = 0 if (score_pat <150 & score_pat !=.)
389 tab score_pa, m
390 drop score_pat miss_plew miss_pleexerc
391
392 foreach var of varlist plew_t pleexerc_t {
393     xtilew `var'3 = `var', within(sex) nq(3)
394 }
395
396 drop w10-dv10 dplew dpleexerc dplesport fdiiy ddiy typepa_a0-typepa_a4 nplew-nplesport
397

```



```

398 *grip strenght*
399 gen gs      = (gsl + gsr)*0.5
400 replace gs = gsl if gsr == .
401 replace gs = gsr if gsl == .
402 mdesc gs*
403 drop gsl gsr
404 *histogram gs
405 *sum gs
406 *replace gs = gs + 1
407 *gen lngs = ln(gs)
408 *histogram lngs
409
410 *walking pace
411 tab wp, m
412 tab wp, m nolab
413 replace wp = . if wp == -7
414 replace wp = . if wp == -3
415 tab wp, m
416
417 saveold "db2", version(13) replace
418 display "$S_TIME  $S_DATE"
419
420 /*further cleaning & define CRF & healthy score diet*/
421 cls
422 cd "Analysis"
423 use "db2", clear
424 describe, short
425
426 *score processed meat
427 tab pmeat, m
428 tab pmeat, m nolab
429 gen score_pm = 1 if (pmeat == 0 | pmeat == 1 | pmeat == 2)
430 replace score_pm = 0 if (pmeat == 3 | pmeat == 4 | pmeat == 5)
431 tab score_pm, m
432
433 *diet variables
434 rename poultry wmeat
435 foreach var of varlist wmeat beef lamb pork pmeat {
436     tab `var', m
437     tab `var', m nolab
438     replace `var' = .    if (`var' == -1 | `var' == -3)
439     gen    n`var' = 0    if `var' == 0
440     replace n`var' = 0.5 if `var' == 1
441     replace n`var' = 1   if `var' == 2
442     replace n`var' = 3   if `var' == 3
443     replace n`var' = 5.5 if `var' == 4
444     replace n`var' = 7.5 if `var' == 5
445     replace n`var' = .   if `var' == .
446 }
447 gen rmeat = nbeef + nlamb + npork
448 drop wmeat *beef *lamb *pork pmeat

```

```

449 rename nwmeat wmeat
450 rename npmeat pmeat
451 mdesc wmeat rmeat pmeat
452 *score red_meat
453 distplot rmeat
454 gen score_rm = 1 if (rmeat <=3 & rmeat !=.)
455 replace score_rm = 0 if (rmeat >3 & rmeat !=.)
456 tab score_rm, m
457
458 histogram rmeat
459 sum rmeat
460 replace rmeat = rmeat + 1
461 gen lnrmear = ln(rmeat)
462 histogram lnrmear
463
464 histogram wmeat
465 sum wmeat
466 *replace wmeat = wmeat + 1
467 *gen lnwmeat = ln(wmeat)
468 *histogram lnwmeat
469
470 histogram pmeat
471 sum pmeat
472 *replace pmeat = pmeat + 1
473 *gen lnpmeat = ln(pmeat)
474 *histogram lnpmeat
475 graph close _all
476
477 cls
478 foreach var of varlist ffruit dfruit cveg rveg cereal {
479     tab `var', m
480     tab `var', m nolab
481     replace `var' = . if (`var' == -1 | `var' == -3)
482     replace `var' = 0.5 if `var' == -10
483 }
484 sum ffruit dfruit cveg rveg
485 gen fvscore = ffruit + dfruit + cveg + rveg
486 drop *fruit *veg
487 mdesc fvscore
488 *score_fv
489 distplot fvscore
490 gen score_fv = 1 if (fvscore >=5 & fvscore !=.)
491 replace score_fv = 0 if (fvscore <5 & fvscore !=.)
492 tab score_fv, m
493
494 histogram fvscore
495 sum fvscore
496 replace fvscore = fvscore + 1
497 gen lnfv = ln(fvscore)
498 histogram lnfv
499 graph close _all

```

```

500
501 histogram cereal
502 sum cereal
503 replace cereal = cereal + 1
504 gen lncer = ln(cereal)
505 histogram lncer          /*highly skewed*/
506 graph close _all
507
508 /*Overall healthy-lifestyle score*/
509 mdesc score_alc score_fv score_pa score_pm score_rm score_sleep score_tv
510 gen hscore = score_alc + score_fv + score_pa + score_pm + score_rm + score_sleep + score_tv
511 mdesc hscore
512 tab hscore, m
513 drop score_*
514
515 *drop pregnant and prev cancer
516 describe, short
517 tab pregn, sort m
518 drop if pregn == 1      /*Yes*/
519 tab bcancer, sort m
520 drop if bcancer == 2    /*Yes*/
521
522 *complete case for covariables of all models (smoke is covariate/effect modifier)
523 describe, short
524 mdesc age0 sex ethn tws smok ndrugs gs wp
525 egen float nmiss = rowmiss(age0 sex ethn tws smok ndrugs gs wp)
526 tab nmiss, m sort
527 drop if nmiss>0
528 describe, short
529
530 *drop not relevant covariables and final modifications
531 drop byear-nnca guim nap-breadt cerealt pregn-fitness targethr-qual alcs-eversmok w mg100 date0-exit_d nmiss maxwl maxhr fhdiy
dhdiy probacc-wearacc
532
533 tab ndrugs
534 histogram ndrugs
535 *replace ndrugs = ndrugs + 1
536 *gen lndrg = ln(ndrugs)
537 *histogram lndrg
538
539 histogram tws
540 sum tws
541 *gladder tws
542 graph close _all
543 describe, short
544 mdesc
545
546 saveold "db3", replace
547 display "$S_TIME" "$S_DATE"
548
549

```

```

550 *****
551 *****
552 **ANALYSIS - DESCRIPTIVE AND MODEL SPECIFICATIONS**
553
554 /******DESCRIPTIVE******/
555 cls
556 display "$S_TIME  $S_DATE"
557 cd "Analysis"
558 use "db3", clear
559 xtilew gs3 = gs, within(sex) nq(3)
560 gen grip = "sex" + string(sex) + "_" + "gs" + string(gs3)
561 sencode grip, gsort(grip) replace
562 gen pace = "sex" + string(sex) + "_" + "pace" + string(wp)
563 sencode pace, gsort(pace) replace
564 describe, short
565
566 /*values PA variables*/
567 preserve
568 clear
569 save "Results_part1\valuesPA", emptyok replace
570 restore
571 forvalues sex = 0/1 {
572     forvalues q = 1/3 {
573         foreach var of varlist totw totm totv plew_t pleexerc_t {
574             preserve
575             qui summstat `var' if sex == `sex' & `var'3 == `q', clear percentiles(50 25 75)
576             gen sex      = "`sex'"
577             gen tertile = `q'
578             append using "Results_part1\valuesPA"
579             save "Results_part1\valuesPA", replace
580             restore
581         }
582     }
583 }
584 preserve
585 use "Results_part1\valuesPA", replace
586 sort sex var tertile
587 drop unique mean-max label
588 gen order = 1 if var == "totw"
589 replace order = 2 if var == "totm"
590 replace order = 3 if var == "totv"
591 replace order = 4 if var == "plew_t"
592 replace order = 5 if var == "pleexerc_t"
593 replace var = "PA walking" if var == "totw"
594 replace var = "PA moderate" if var == "totm"
595 replace var = "PA vigorous" if var == "totv"
596 replace var = "Pleasure PA walking" if var == "plew_t"
597 replace var = "Pleasure PA exercise" if var == "pleexerc_t"
598 replace sex = "Women" if sex == "0"

```

```

599  replace sex      = "Men"                                if sex == "1"
600  foreach var of varlist p50-p75 {
601      tostring `var', format(%7.0f) replace force
602  }
603  gen min_week = p50 + " (" + p25 + "-" + p75 + ")"
604  gsort -sex order tertile
605  order sex var tertile n min_week
606  keep sex var tertile n min_week
607  export excel using "Results_part1\valuePA.xls", firstrow(variables) replace
608  restore
609
610  /*baseline tables*/
611  preserve
612
613  foreach var of varlist rmeat fvs cer tv {
614      replace `var' = `var' - 1                                /*changes above for ln transformation*/
615  }
616
617  baselinetable                                                /*
618  */ age0(cts tab("p50 (p25-p75)"))                            /*
619  */ ethn(cat)                                                  /*
620  */ tws(cts tab("p50 (p25-p75)"))                            /*
621  */ smok(cat)                                                  /*
622  */ ndrugs(cts tab("p50 (p25-p75)"))                          /*
623  */ bmi(cts tab("p50 (p25-p75)"))                            /*
624  */ alcf(cat)                                                  /*
625  */ wmeat(cts tab("p50 (p25-p75)"))                          /*
626  */ rmeat(cts tab("p50 (p25-p75)"))                          /*
627  */ pmeat(cts tab("p50 (p25-p75)"))                          /*
628  */ fvscore(cts tab("p50 (p25-p75)"))                        /*
629  */ cereal(cts tab("p50 (p25-p75)"))                        /*
630  */ tv(cts tab("p50 (p25-p75)"))                            /*
631  */ cpu(cts tab("p50 (p25-p75)"))                            /*
632  */ sleep(cts tab("p50 (p25-p75)"))                          /*
633  */ totw3(cat)                                                /*
634  */ totm3(cat)                                                /*
635  */ totv3(cat)                                                /*
636  */ plew_t3(cat)                                              /*
637  */ pleexerc_t3(cat)                                          /*
638  */ hscore(cts tab("p50 (p25-p75)"))                        /*
639  */ wp(cat)                                                  /*
640  */ gs(cts tab("p50 (p25-p75)"))                            /*
641  */ bcvd(cat)                                                 /*
642  */ bcva(cat)                                                 /*
643  */ bckd(cat)                                                 /*
644  */ all_death(cat)                                            /*
645  */ , reportmissing by(sex) notable exportexcel("Results_part1\table_1", replace)
646
647  baselinetable                                                /*
648  */ age0(cts tab("p50 (p25-p75)"))                            /*
649  */ ethn(cat)                                                  /*

```

```

650 */ tws(cts tab("p50 (p25-p75)")) /*
651 */ smok(cat) /*
652 */ ndrugs(cts tab("p50 (p25-p75)")) /*
653 */ bmi(cts tab("p50 (p25-p75)")) /*
654 */ alcf(cat) /*
655 */ wmeat(cts tab("p50 (p25-p75)")) /*
656 */ rmeat(cts tab("p50 (p25-p75)")) /*
657 */ pmeat(cts tab("p50 (p25-p75)")) /*
658 */ fvscore(cts tab("p50 (p25-p75)")) /*
659 */ cereal(cts tab("p50 (p25-p75)")) /*
660 */ tv(cts tab("p50 (p25-p75)")) /*
661 */ cpu(cts tab("p50 (p25-p75)")) /*
662 */ sleep(cts tab("p50 (p25-p75)")) /*
663 */ totw3(cat) /*
664 */ totm3(cat) /*
665 */ totv3(cat) /*
666 */ plew_t3(cat) /*
667 */ pleexerc_t3(cat) /*
668 */ hscore(cts tab("p50 (p25-p75)")) /*
669 */ wp(cat) /*
670 */ gs(cts tab("p50 (p25-p75)")) /*
671 */ bcvd(cat) /*
672 */ bcva(cat) /*
673 */ bckd(cat) /*
674 */ all_death(cat) /*
675 */ , reportmissing by(grip) notable exportexcel("Results_part1\table_s1", replace)
676
677 baselinetable /*
678 */ age0(cts tab("p50 (p25-p75)")) /*
679 */ ethn(cat) /*
680 */ tws(cts tab("p50 (p25-p75)")) /*
681 */ smok(cat) /*
682 */ ndrugs(cts tab("p50 (p25-p75)")) /*
683 */ bmi(cts tab("p50 (p25-p75)")) /*
684 */ alcf(cat) /*
685 */ wmeat(cts tab("p50 (p25-p75)")) /*
686 */ rmeat(cts tab("p50 (p25-p75)")) /*
687 */ pmeat(cts tab("p50 (p25-p75)")) /*
688 */ fvscore(cts tab("p50 (p25-p75)")) /*
689 */ cereal(cts tab("p50 (p25-p75)")) /*
690 */ tv(cts tab("p50 (p25-p75)")) /*
691 */ cpu(cts tab("p50 (p25-p75)")) /*
692 */ sleep(cts tab("p50 (p25-p75)")) /*
693 */ totw3(cat) /*
694 */ totm3(cat) /*
695 */ totv3(cat) /*
696 */ plew_t3(cat) /*
697 */ pleexerc_t3(cat) /*
698 */ hscore(cts tab("p50 (p25-p75)")) /*
699 */ wp(cat) /*
700 */ gs(cts tab("p50 (p25-p75)")) /*

```

```

701 */ bcvd(cat) */
702 */ bcva(cat) */
703 */ bckd(cat) */
704 */ all_death(cat) */
705 */ , reportmissing by(pace) notable exportexcel("Results_part1\table_s2", replace)
706
707 display "$S_TIME $S_DATE"
708 restore
709 drop bmi totw totm totv plew_t pleexerc_t rmeat fvscore cereal tv
710
711 /#####*****SURVIVAL*****#####*/
712 *###RATES
713 describe, short
714 preserve
715 clear
716 tempfile rates
717 save `rates', emptyok replace
718 restore
719 foreach out in all_death {
720     foreach exp in grip pace {
721         preserve
722         qui parmby "poisson `out', exp(time) irr", by(`exp') es(N) fast
723         append using `rates'
724         save `rates', replace
725         restore
726     }
727 }
728 preserve
729 use `rates', clear
730 sdecode pace, replace
731 sdecode grip, replace
732 gen group = pace if grip == ""
733 replace group = grip if pace == ""
734 gen rate1000 = exp(est)*1000
735 gen lb95 = exp(min95)*1000
736 gen ub95 = exp(max95)*1000
737 keep eq group rate1000 lb ub es_1
738 rename es_1 Npart
739 rename eq outcome
740 foreach var of varlist rate1000-ub95 {
741     tostring `var', format(%7.2f) replace force
742 }
743 gen rate = rate1000 + " (" + lb95 + ", " + ub95 + ")"
744 export excel using "Results_part1\rates.xls", firstrow(variables) replace
745 display "$S_TIME $S_DATE"
746 restore
747
748 *###MODEL SPECIFICATION
749 /*assess AIC and LRtest for linearity and interaction, both sex, all-cause death*/
750 /*Principle: for categories, see if interaction, then use cat##modifier or cat + modifier according to LRtest*/
751 /*for continuous, used spline interaction and compared with spline no interaction, then used rcs(lin)##modifier or rcs(lin) +

```

```
modifier according to LRtest*/
```

```

752
753 stset time, id(n_eid) failure(all_death==1)
754 strate, per(1000) /*overall rate*/
755 stdescribe
756 stsum
757
758 preserve
759 clear
760 tempfile models
761 gen sex = .
762 gen exp = .
763 gen mod = .
764 gen type = .
765 gen aic = .
766 gen aic_i = .
767 gen p_int = .
768 save `models', emptyok replace
769 restore
770 forvalues k = 0/1 {
771     foreach exp in gs3 wp {
772         foreach var of varlist smok alcf totw3 totm3 totv3 plew_t3 pleexerc_t3 /*cat*/ {
773             preserve
774             cap nois stpm2 age0 ethn tws i.smok ndrugs i.`exp' i.`var' if sex == `k', scale(hazard) df(4) eform lininit
775             /*ignored by stata if repeated exposure smok*/
776             estat ic
777             mat s=r(S)
778             local cat_aic = s[1,6]
779             estimate store nointer
780             cap nois stpm2 age0 ethn tws i.smok ndrugs i.`exp'##i.`var' if sex == `k', scale(hazard) df(4) eform lininit
781             /*ignored by stata if repeated exposure smok*/
782             estat ic
783             mat s=r(S)
784             local cati_aic = s[1,6]
785             estimate store inter
786             lrtest (nointer) (inter)
787             local p = r(p)
788             clear
789             set obs 1
790             gen sex = "`k'"
791             gen exp = "`exp'"
792             gen mod = "`var'"
793             gen type = "cat"
794             gen aic = `cat_aic'
795             gen aic_i = `cati_aic'
796             gen p_int = `p'
797             append using `models', force
798             save `models', replace
799             restore
800         }
801     }
802 }

```



```

800 }
801 foreach var of varlist wmeat lnrmeat pmeat lnfvb lncr lntv cpu sleep lnbmi hscore {
802     rcsgen `var', df(3) gen(spl_`var') orthog
803 }
804 forvalues k = 0/1 {
805     foreach exp in gs3 wp {
806         foreach var of varlist wmeat lnrmeat pmeat lnfvb lncr lntv cpu sleep lnbmi hscore /*cont*/ {
807             preserve
808             cap nois stpm2 age0 ethn tws i.smok ndrugs i.`exp' c.spl_`var'* if sex == `k', scale(hazard) df(4) eform lininit
809             estat ic
810             mat s=r(S)
811             local lin_aic = s[1,6]
812             estimate store nointer
813             cap nois stpm2 age0 ethn tws i.smok ndrugs i.`exp'##c.spl_`var'* if sex == `k', scale(hazard) df(4) eform lininit
814             estat ic
815             mat s=r(S)
816             local lini_aic = s[1,6]
817             estimate store inter
818             lrtest (nointer) (inter)
819             local p = r(p)
820             clear
821             set obs 1
822             gen sex = "`k'"
823             gen exp = "`exp'"
824             gen mod = "`var'"
825             gen type = "cont_spline"
826             gen aic = `lin_aic'
827             gen aic_i = `lini_aic'
828             gen p_int = `p'
829             append using `models', force
830             save `models', replace
831             restore
832         }
833     }
834 }
835 preserve
836 use `models', clear
837 gen out = "all_death"
838 sort sex exp type
839 gen cutoff = 0.05/17 /*number of tests within exposure, gs3 wp*/
840 gen sig = 1 if p_int < cutoff
841 gen dif = aic-aic_i if sig == 1 /*better no interaction; no interaction to be used, with spline for continuous*/
842 gen dif2 = aic-aic_i /*all differences*/
843 sort dif2
844 tostring sex, replace
845 replace sex = "Women" if sex == "0"
846 replace sex = "Men" if sex == "1"
847 export excel using "Results_part1\interactions.xls", firstrow(variables) replace
848 display "$S_TIME $S_DATE"
849 restore
850

```

```

851  ***DISTRIBUTION COVARIATES FOR PREDICTION
852  foreach var of varlist wmeat lnrmear pmeat lnfvvs lncr lntv cpu sleep lnbmi hscore {
853      distplot `var', name(`var', replace) nodraw
854  }
855  graph combine wmeat lnrmear pmeat lnfvvs lncr lntv cpu sleep lnbmi hscore
856  graph close _all
857
858  baselinetable                                /*
859  /* wmeat(cts tab("p50 (p5-p95)"))          /*
860  /* lnrmear(cts tab("p50 (p5-p95)"))          /*
861  /* pmeat(cts tab("p50 (p5-p95)"))          /*
862  /* lnfvvs(cts tab("p50 (p5-p95)"))          /*
863  /* lncr(cts tab("p50 (p5-p95)"))          /*
864  /* lntv(cts tab("p50 (p5-p95)"))          /*
865  /* cpu(cts tab("p50 (p5-p95)"))          /*
866  /* sleep(cts tab("p50 (p5-p95)"))          /*
867  /* lnbmi(cts tab("p50 (p5-p95)"))          /*
868  /* hscore(cts tab("p50 (p5-p95)"))          /*
869  /* , by(sex) notable exportexcel("Results_part1\table_s3", replace) medianformat(%5.2f)
870  display "$S_TIME  $S_DATE"
871
872
873
874  #####
875  #####
876  **FLEXIBLE SURVIVAL REGRESSIONS - SURVIVAL CONTINUOUS MODIFIERS**
877  cls
878  cd "Analysis"
879  use "db3", clear
880  xtilew gs3 = gs, within(sex) nq(3)
881  drop bmi totw totm totv plew_t pleexerc_t rmeat fvscore cereal tv
882  rename gs3 gt                                /*make easier the loop*/
883  foreach var of varlist gt wp smok {
884      tab `var', gen(`var')
885  }
886  gen timevar = 10 in 1
887
888  //////////////////////////////////////
889  cls
890  local lev_wmeat 0(0.15)3
891  local lev_lnrmear 0(0.0895)1.79
892  local lev_pmeat 0(0.275)5.5
893  local lev_lnfvs 1.1(0.0865)2.83
894  local lev_lncr 0(0.104)2.08
895  local lev_lntv 0.41(0.077)1.95
896  local lev_cpu 0(0.2)4
897  local lev_sleep 5(0.2)9
898  local lev_lnbmi 3.02(0.0295)3.61
899  local lev_hscore 2(0.25)7

```

```

900
901 tempfile survs_co
902 cap postclose stats
903 postfile stats str10 sex str10 out str10 exp str25 mod npart events s1 lb1 ub1 s2 lb2 ub2 s3 lb3 ub3 d3 lbd3 ubd3 xval using
   `survs_co'
904
905 forvalues sex = 0/1 {
906     foreach mod of varlist wmeat lnrmear pmeat lnfv s lncr
907         /*                               lntv cpu sleep hscore lnbmi {                               /* this is the logical order for later graph */
908             cap drop `mod's*
909             rcsgen `mod', gen(`mod's) orthog df(3)
910             global K`mod' `r(knots)'
911             matrix M`mod' = r(R)
912             foreach out in all_death {
913                 stset time, id(n_eid) failure(`out'==1)
914                 tabstat `out' if (`mod' !=. & sex == `sex'), statistics(count sum) save
915                 qui tabstatmat mx
916                 local npart = mx[1,1]
917                 local events = mx[2,1]
918                 foreach exp in gt wp {
919                     cap nois stpm2 age0 ethn tws smok2 smok3 ndrugs `exp'2 `exp'3 `mod's1 `mod's2 `mod's3 if sex == `sex', scale(hazard
920 ) df(4) eform lininit
921                     forvalues k = `lev `mod'' {
922                         cap drop surv1* surv2* surv3* d2* d3*
923                         rcsgen, scalar(`k') knots(`${K`mod}`) rmatrix(M`mod') gen(v)
924                         stpm2_standsurv, atvars(surv1 surv2 surv3)
925                         /* at1("`exp'2" 0 "`exp'3" 0 "`mod's1" `=v1' "`mod's2" `=v2' "`mod's3" `=v3') /*
926                         /* at2("`exp'2" 1 "`exp'3" 0 "`mod's1" `=v1' "`mod's2" `=v2' "`mod's3" `=v3') /*
927                         /* at3("`exp'2" 0 "`exp'3" 1 "`mod's1" `=v1' "`mod's2" `=v2' "`mod's3" `=v3') /*
928                         /* timevar(timevar) ci contrast(difference) contrastvars(d2 d3)
929                         local s1 = surv1[1]
930                         local s2 = surv2[1]
931                         local s3 = surv3[1]
932                         local lb1 = surv1_lci[1]
933                         local lb2 = surv2_lci[1]
934                         local lb3 = surv3_lci[1]
935                         local ub1 = surv1_uci[1]
936                         local ub2 = surv2_uci[1]
937                         local ub3 = surv3_uci[1]
938                         local d3 = d3[1]
939                         local lbd3 = d3_lci[1]
940                         local ubd3 = d3_uci[1]
941                         local xval = `k'
942                         post stats (`sex'" ("`out'" ("`exp'" ("`mod'" ("`npart'" (`events'" (`s1'" (`lb1'" (`ub1'" (`s2'" (`lb2'" (
943                         `ub2'" (`s3'" (`lb3'" (`ub3'" (`d3'" (`lbd3'" (`ubd3'" (`xval'"
944                     }
945                 }
946             }
947 postclose stats

```

```

948
949 *for BMI there is interaction with walking pace in men, so results need to be changed*
950 cls
951 cd "Analysis"
952 use "db3", clear
953 xtilew gs3 = gs, within(sex) nq(3)
954 drop bmi totw totm totv plew_t pleexerc_t rmeat fvscore cereal tv
955 foreach var of varlist wp smok {
956     tab `var', gen(`var')
957 }
958 keep if sex == 1                                /*interaction only in men*/
959 gen timevar = 10 in 1
960
961 rcsgen lnbmi, gen(lnbmis) orthog df(3)
962 global Klnbmi `r(knots)'
963 matrix Mlnbmi = r(R)
964
965 order lnbmi, before(lnbmis1)
966 foreach mod of varlist lnbmis* {
967     foreach exp of varlist wp2-wp3 {
968         gen `mod'_'exp' = `mod'*`exp'
969     }
970 }
971
972 tempfile survs_cobmi
973 cap postclose statsbmi
974 postfile statsbmi str10 sex str10 out str10 exp str25 mod npart events s1 lb1 ub1 s2 lb2 ub2 s3 lb3 ub3 d3 lbd3 ubd3 xval using
`survs_cobmi'
975
976 stset time, id(n_eid) failure(all_death==1)
977 tabstat all_death if lnbmi !=., statistics(count sum) save
978 qui tabstatmat mx
979 local npart = mx[1,1]
980 local events = mx[2,1]
981
982 stpm2 age0 ethn tws smok2 smok3 ndrugs wp2 wp3 lnbmis1 lnbmis2 lnbmis3 lnbmis1_wp2-lnbmis3_wp3, scale(hazard) df(4) eform lininit
983
984 local lev_lnbmi 3.02(0.0295)3.61
985
986 forvalues k = `lev_lnbmi' {
987     cap drop surv1* surv2* surv3* d2* d3*
988     rcsgen, scalar(`k') knots(`${Klnbmi}`) rmatrix(Mlnbmi) gen(v)
989     stpm2_standsurv, atvars(surv1 surv2 surv3) /*
990     /* at1(wp2 0 wp3 0 lnbmis1 `=v1' lnbmis2 `=v2' lnbmis3 `=v3' lnbmis1_wp2 0 lnbmis1_wp3 0 lnbmis2_wp2 0
lnbmis2_wp3 0 lnbmis3_wp2 0 lnbmis3_wp3 0) /*
991     /* at2(wp2 1 wp3 0 lnbmis1 `=v1' lnbmis2 `=v2' lnbmis3 `=v3' lnbmis1_wp2 `=v1' lnbmis1_wp3 0 lnbmis2_wp2 `=v2'
lnbmis2_wp3 0 lnbmis3_wp2 `=v3' lnbmis3_wp3 0) /*
992     /* at3(wp2 0 wp3 1 lnbmis1 `=v1' lnbmis2 `=v2' lnbmis3 `=v3' lnbmis1_wp2 0 lnbmis1_wp3 `=v1' lnbmis2_wp2 0
lnbmis2_wp3 `=v2' lnbmis3_wp2 0 lnbmis3_wp3 `=v3') /*
993     /* timevar(timevar) ci contrast(difference) contrastvars(d2 d3)
994     local s1 = surv1[1]

```

```

995     local s2 = surv2[1]
996     local s3 = surv3[1]
997     local lb1 = surv1_lci[1]
998     local lb2 = surv2_lci[1]
999     local lb3 = surv3_lci[1]
1000    local ub1 = surv1_uci[1]
1001    local ub2 = surv2_uci[1]
1002    local ub3 = surv3_uci[1]
1003    local d3 = d3[1]
1004    local lbd3 = d3_lci[1]
1005    local ubd3 = d3_uci[1]
1006    local xval = `k'
1007    post statsbmi("`sex'"')("`out'"')("`exp'"')("`mod'"')(`npart')(`events')(`s1')(`lb1')(`ub1')(`s2')(`lb2')(`ub2')(`s3')(`lb3')(`ub3')(`d3')(`lbd3')(`ubd3')(`xval')
1008  }
1009  postclose statsbmi
1010
1011  *combine results*
1012  use `survs_co', clear
1013  drop if (mod == "lnbmi" & exp == "wp" & sex == "1") /*from analysis without interaction*/
1014  append using `survs_cobmi', force /*from analysis with interaction*/
1015  replace sex = "1" if sex == ""
1016  replace out = "all death" if out == ""
1017  replace exp = "wp" if exp == ""
1018  replace mod = "lnbmi" if mod == ""
1019  replace sex = "Women" if sex == "0"
1020  replace sex = "Men" if sex == "1"
1021  foreach var of varlist sex-mod {
1022    sencode `var', replace
1023  }
1024  bys sex: tab exp mod
1025  display "$S_TIME $S_DATE"
1026  save "Results_part1\survs_co", replace
1027
1028
1029  #####
1030  #####
1031  **FLEXIBLE SURVIVAL REGRESSIONS - SURVIVAL CATEGORICAL MODIFIERS**
1032  cls
1033  cd "Analysis"
1034  use "db3", clear
1035  xtilew gs3 = gs, within(sex) nq(3)
1036  drop bmi totw totm totv plew_t pleexerc_t rmeat fvscore cereal tv
1037  rename gs3 gt /*make easier the loop*/
1038  rename totw3 totw
1039  rename totm3 totm
1040  rename totv3 totv
1041  rename plew_t3 plew_t
1042  rename pleexerc_t3 pleexerc_t

```

```

1043 foreach var of varlist gt wp {
1044     tab `var', gen(`var')
1045 }
1046 foreach var of varlist smok alcf totw totm totv plew_t pleexerc_t {
1047     tab `var', gen(`var')
1048 }
1049 drop alcf smok totw totm totv plew_t pleexerc_t alcf1 smok1 totw1 totm1 totv1 plew_t1 pleexerc_t1    /*to facilitate wildcard stpm2
below in the loop*/
1050 gen timevar = 10 in 1
1051
1052
1053 *PHYSICAL ACTIVITY VARIABLES - TERTILES//////////////////////////////////////
1054 cls
1055 tempfile surv_pa /*3 levels/tertiles */
1056 cap postclose stats
1057 postfile stats str10 sex str10 out str10 exp str15 mod npart events /*
1058 */ s11 lb11 ub11 s12 lb12 ub12 s13 lb13 ub13 /*
1059 */ s21 lb21 ub21 s22 lb22 ub22 s23 lb23 ub23 /*
1060 */ s31 lb31 ub31 s32 lb32 ub32 s33 lb33 ub33 /*
1061 */ using `surv_pa'
1062
1063 preserve
1064 clear
1065 tempfile difpa
1066 gen sex      = .
1067 gen exp      = .
1068 gen dif      = .
1069 gen lb_dif   = .
1070 gen ub_dif   = .
1071 gen contr    = .
1072 gen mod      = .
1073 save `difpa', emptyok replace
1074 restore
1075
1076 mata: mata clear
1077 mata:
1078     function d1(at) {
1079         return(at[7] - at[1])    /*contrasts third tertile vs first, level mod 1 */
1080     }
1081 end
1082 mata:
1083     function d2(at) {
1084         return(at[8] - at[2])    /*contrasts third tertile vs first, level mod 2*/
1085     }
1086 end
1087 mata:
1088     function d3(at) {
1089         return(at[9] - at[3])    /*contrasts third tertile vs first, level mod 3*/
1090     }
1091 end
1092

```

```

1093   forvalues sex = 0/1 {
1094       foreach mod in totw totm totv plew_t pleexerc_t {
1095           foreach out in all_death {
1096               cap drop `mod's*
1097               tabstat `out' if (`mod'2 !=. & sex == `sex'), statistics(count sum) save
1098               qui tabstatmat mx
1099               local npart = mx[1,1]
1100               local events = mx[2,1]
1101               stset time, id(n_eid) failure(`out'==1)
1102               foreach exp in gt wp {
1103                   stpm2 age0 ethn tws smok2 smok3 ndrugs `exp'2 `exp'3 `mod'2 `mod'3 if sex == `sex', scale(hazard) df(4) eform
1104                   forvalues m = 1/3 {
1105                       cap drop s1* s2* s3* uf*
1106                       stpm2_standsurv, atvars(s11 s12 s13 s21 s22 s23 s31 s32 s33) /*
1107                       /* at1("`exp'2" 0 "`exp'3" 0 `mod'2 0 `mod'3 0) /*
1108                       /* at2("`exp'2" 0 "`exp'3" 0 `mod'2 1 `mod'3 0) /*
1109                       /* at3("`exp'2" 0 "`exp'3" 0 `mod'2 0 `mod'3 1) /*
1110                       /* at4("`exp'2" 1 "`exp'3" 0 `mod'2 0 `mod'3 0) /*
1111                       /* at5("`exp'2" 1 "`exp'3" 0 `mod'2 1 `mod'3 0) /*
1112                       /* at6("`exp'2" 1 "`exp'3" 0 `mod'2 0 `mod'3 1) /*
1113                       /* at7("`exp'2" 0 "`exp'3" 1 `mod'2 0 `mod'3 0) /*
1114                       /* at8("`exp'2" 0 "`exp'3" 1 `mod'2 1 `mod'3 0) /*
1115                       /* at9("`exp'2" 0 "`exp'3" 1 `mod'2 0 `mod'3 1) /*
1116                       /* timevar(timevar) ci userfunction(d`m') userfunctionvar(uf)
1117                       local j = `m' + 6
1118                       local dif = uf[1]
1119                       local lb_dif = uf_lci[1]
1120                       local ub_dif = uf_uci[1]
1121                       preserve
1122                       clear
1123                       set obs 1
1124                       gen sex = "`sex'"
1125                       gen exp = "`exp'"
1126                       gen dif = `dif'
1127                       gen lb_dif = `lb_dif'
1128                       gen ub_dif = `ub_dif'
1129                       gen contr = "ct_`j'`m'"
1130                       gen mod = "`mod'"
1131                       append using `difpa', force
1132                       save `difpa', replace
1133                       restore
1134                   }
1135                   foreach t of numlist 11/13 21/23 31/33 {
1136                       local s`t' = s`t'[1]
1137                       local lb`t' = s`t'_lci[1]
1138                       local ub`t' = s`t'_uci[1]
1139                   }
1140                   post stats (`sex') (`out') (`exp') (`mod') (`npart') (`events') /*
1141                   /* (`s11') (`lb11') (`ub11') (`s12') (`lb12') (`ub12') (`s13') (`lb13') (`ub13') /*
1142                   /* (`s21') (`lb21') (`ub21') (`s22') (`lb22') (`ub22') (`s23') (`lb23') (`ub23') /*
1143                   /* (`s31') (`lb31') (`ub31') (`s32') (`lb32') (`ub32') (`s33') (`lb33') (`ub33')

```

```

1144     }
1145 }
1146 }
1147 }
1148
1149 postclose stats
1150 preserve
1151 use `surv_pa', clear
1152
1153 gen i = _n
1154 reshape long s lb ub, i(i) j(group)
1155 drop i
1156 gen level_exp = "First" if inlist(group, 11, 12, 13)
1157 replace level_exp = "Second" if inlist(group, 21, 22, 23)
1158 replace level_exp = "Third" if inlist(group, 31, 32, 33)
1159 gen level_mod = "First Tertile" if inlist(group, 11, 21, 31)
1160 replace level_mod = "Second Tertile" if inlist(group, 12, 22, 32)
1161 replace level_mod = "Third Tertile" if inlist(group, 13, 23, 33)
1162 order sex out exp level_e mod level_m s lb ub np ev group
1163 save "Results_part1\surv_pa", replace
1164
1165 use `difpa', clear
1166 gen level_mod = "First Tertile" if contr == "ct_71"
1167 replace level_mod = "Second Tertile" if contr == "ct_82"
1168 replace level_mod = "Third Tertile" if contr == "ct_93"
1169 drop contr
1170 save "Results_part1\survdif_pa", replace
1171 display "$S_TIME $S_DATE"
1172 restore
1173
1174
1175 *SMOKING//////////////////////////////////////
1176 cls
1177 tempfile surv_smok /*3 levels*/
1178 cap postclose stats
1179 postfile stats str10 sex str10 out str10 exp npart events /*
1180 */ s11 lb11 ub11 s12 lb12 ub12 s13 lb13 ub13 /*
1181 */ s21 lb21 ub21 s22 lb22 ub22 s23 lb23 ub23 /*
1182 */ s31 lb31 ub31 s32 lb32 ub32 s33 lb33 ub33 /*
1183 */ using `surv_smok'
1184
1185 preserve
1186 clear
1187 tempfile difsmok
1188 gen sex = .
1189 gen exp = .
1190 gen dif = .
1191 gen lb_dif = .
1192 gen ub_dif = .
1193 gen contr = .
1194 gen mod = .

```



```

1195 save `difsmok', emptyok replace
1196 restore
1197
1198 mata: mata clear
1199 mata:
1200     function d1(at) {
1201         return(at[7] - at[1])    /*contrasts third tertile vs first, level mod 1 */
1202     }
1203 end
1204 mata:
1205     function d2(at) {
1206         return(at[8] - at[2])    /*contrasts third tertile vs first, level mod 2*/
1207     }
1208 end
1209 mata:
1210     function d3(at) {
1211         return(at[9] - at[3])    /*contrasts third tertile vs first, level mod 3*/
1212     }
1213 end
1214
1215 forvalues sex = 0/1 {
1216     foreach out in all_death {
1217         tabstat `out' if (smok2 !=. & sex == `sex'), statistics(count sum) save
1218         qui tabstatmat mx
1219         local npart = mx[1,1]
1220         local events = mx[2,1]
1221         stset time, id(n_eid) failure(`out'==1)
1222         foreach exp in gt wp {
1223             stpm2 age0 ethn tws smok2 smok3 ndrugs `exp'2 `exp'3 if sex == `sex', scale(hazard) df(4) eform
1224             forvalues m = 1/3 {
1225                 cap drop s1* s2* s3* uf*
1226                 stpm2_standsurv, atvars(s11 s12 s13 s21 s22 s23 s31 s32 s33) /*
1227                 /* at1("`exp'2" 0 "`exp'3" 0 smok2 0 smok3 0) /*
1228                 /* at2("`exp'2" 0 "`exp'3" 0 smok2 1 smok3 0) /*
1229                 /* at3("`exp'2" 0 "`exp'3" 0 smok2 0 smok3 1) /*
1230                 /* at4("`exp'2" 1 "`exp'3" 0 smok2 0 smok3 0) /*
1231                 /* at5("`exp'2" 1 "`exp'3" 0 smok2 1 smok3 0) /*
1232                 /* at6("`exp'2" 1 "`exp'3" 0 smok2 0 smok3 1) /*
1233                 /* at7("`exp'2" 0 "`exp'3" 1 smok2 0 smok3 0) /*
1234                 /* at8("`exp'2" 0 "`exp'3" 1 smok2 1 smok3 0) /*
1235                 /* at9("`exp'2" 0 "`exp'3" 1 smok2 0 smok3 1) /*
1236                 /* timevar(timevar) ci userfunction(d`m') userfunctionvar(uf)
1237                 local j = `m' + 6
1238                 local dif = uf[1]
1239                 local lb_dif = uf_lci[1]
1240                 local ub_dif = uf_uci[1]
1241                 preserve
1242                 clear
1243                 set obs 1
1244                 gen sex = "`sex'"
1245                 gen exp = "`exp'"

```

```

1246         gen dif      = `dif'
1247         gen lb_dif    = `lb_dif'
1248         gen ub_dif    = `ub_dif'
1249         gen contr     = "ct_`j'`m'"
1250         gen mod       = "Smoking"
1251         append using `difsmok', force
1252         save `difsmok', replace
1253         restore
1254     }
1255     foreach t of numlist 11/13 21/23 31/33 {
1256         local s`t' = s`t'[1]
1257         local lb`t' = s`t' _lci[1]
1258         local ub`t' = s`t' _uci[1]
1259     }
1260     post stats  ("`sex'"') ("`out'"') ("`exp'"') (`npart') (`events') /*
1261     */          (`s11') (`lb11') (`ub11') (`s12') (`lb12') (`ub12') (`s13') (`lb13') (`ub13') /*
1262     */          (`s21') (`lb21') (`ub21') (`s22') (`lb22') (`ub22') (`s23') (`lb23') (`ub23') /*
1263     */          (`s31') (`lb31') (`ub31') (`s32') (`lb32') (`ub32') (`s33') (`lb33') (`ub33')
1264 }
1265 }
1266 }
1267 postclose stats
1268 preserve
1269 use `surv_smok', clear
1270 gen i = _n
1271 reshape long s lb ub, i(i) j(group)
1272 drop i
1273 gen level_exp      = "First"    if inlist(group, 11, 12, 13)
1274 replace level_exp  = "Second"   if inlist(group, 21, 22, 23)
1275 replace level_exp  = "Third"    if inlist(group, 31, 32, 33)
1276 gen level_mod      = "Never"    if inlist(group, 11, 21, 31)
1277 replace level_mod  = "Former"   if inlist(group, 12, 22, 32)
1278 replace level_mod  = "Current"  if inlist(group, 13, 23, 33)
1279 gen mod = "Smoking"
1280 order sex out exp level_e mod level_m s lb ub np ev group
1281 save "Results_part1\surv_smok", replace
1282
1283 use `difsmok', clear
1284 gen level_mod      = "Never"    if contr == "ct_71"
1285 replace level_mod  = "Former"   if contr == "ct_82"
1286 replace level_mod  = "Current"  if contr == "ct_93"
1287 drop contr
1288 save "Results_part1\survdiff_smok", replace
1289 display "$S_TIME  $S_DATE"
1290 restore
1291
1292 *ALCOHOL FREQUENCY//////////////////////////////////////
1293 cls
1294 tempfile surv_alc      /*6 levels*/
1295 cap postclose stats
1296 postfile stats strl0 sex strl0 out strl0 exp npart events /*

```

```

1297  */ s11 lb11 ub11 s12 lb12 ub12 s13 lb13 ub13 s14 lb14 ub14 s15 lb15 ub15 s16 lb16 ub16 /*
1298  */ s21 lb21 ub21 s22 lb22 ub22 s23 lb23 ub23 s24 lb24 ub24 s25 lb25 ub25 s26 lb26 ub26 /*
1299  */ s31 lb31 ub31 s32 lb32 ub32 s33 lb33 ub33 s34 lb34 ub34 s35 lb35 ub35 s36 lb36 ub36 /*
1300  */ using `surv_alc'
1301
1302  preserve
1303  clear
1304  tempfile difalc
1305  gen sex      = .
1306  gen exp      = .
1307  gen dif      = .
1308  gen lb_dif   = .
1309  gen ub_dif   = .
1310  gen contr    = .
1311  gen mod      = .
1312  save `difalc', emptyok replace
1313  restore
1314
1315  mata: mata clear
1316  mata:
1317      function d1(at) {
1318          return(at[13] - at[1]) /*contrasts third tertile vs first, level mod 1 */
1319      }
1320  end
1321  mata:
1322      function d2(at) {
1323          return(at[14] - at[2]) /*contrasts third tertile vs first, level mod 2*/
1324      }
1325  end
1326  mata:
1327      function d3(at) {
1328          return(at[15] - at[3]) /*contrasts third tertile vs first, level mod 3*/
1329      }
1330  end
1331  mata:
1332      function d4(at) {
1333          return(at[16] - at[4]) /*contrasts third tertile vs first, level mod 4*/
1334      }
1335  end
1336  mata:
1337      function d5(at) {
1338          return(at[17] - at[5]) /*contrasts third tertile vs first, level mod 5*/
1339      }
1340  end
1341  mata:
1342      function d6(at) {
1343          return(at[18] - at[6]) /*contrasts third tertile vs first, level mod 6*/
1344      }
1345  end
1346
1347  forvalues sex = 0/1 {

```

```

1348   foreach out in all_death {
1349       tabstat `out' if (alcf2 !=. & sex == `sex'), statistics(count sum) save
1350       qui tabstatmat mx
1351       local npart = mx[1,1]
1352       local events = mx[2,1]
1353       stset time, id(n_eid) failure(`out'==1)
1354       foreach exp in gt wp {
1355           stpm2 age0 ethn tws smok2 smok3 ndrugs `exp'2 `exp'3 alcf* if sex == `sex', scale(hazard) df(4) eform
1356           forvalues m = 1/6 {
1357               cap drop s1* s2* s3* uf*
1358               stpm2 standsurv, atvars(s11 s12 s13 s14 s15 s16 s21 s22 s23 s24 s25 s26 s31 s32 s33 s34 s35 s36) /*
1359               /* at1( "`exp'2" 0 "`exp'3" 0 alcf2 0 alcf3 0 alcf4 0 alcf5 0 alcf6 0) /*
1360               /* at2( "`exp'2" 0 "`exp'3" 0 alcf2 1 alcf3 0 alcf4 0 alcf5 0 alcf6 0) /*
1361               /* at3( "`exp'2" 0 "`exp'3" 0 alcf2 0 alcf3 1 alcf4 0 alcf5 0 alcf6 0) /*
1362               /* at4( "`exp'2" 0 "`exp'3" 0 alcf2 0 alcf3 0 alcf4 1 alcf5 0 alcf6 0) /*
1363               /* at5( "`exp'2" 0 "`exp'3" 0 alcf2 0 alcf3 0 alcf4 0 alcf5 1 alcf6 0) /*
1364               /* at6( "`exp'2" 0 "`exp'3" 0 alcf2 0 alcf3 0 alcf4 0 alcf5 0 alcf6 1) /*
1365               /* at7( "`exp'2" 1 "`exp'3" 0 alcf2 0 alcf3 0 alcf4 0 alcf5 0 alcf6 0) /*
1366               /* at8( "`exp'2" 1 "`exp'3" 0 alcf2 1 alcf3 0 alcf4 0 alcf5 0 alcf6 0) /*
1367               /* at9( "`exp'2" 1 "`exp'3" 0 alcf2 0 alcf3 1 alcf4 0 alcf5 0 alcf6 0) /*
1368               /* at10("`exp'2" 1 "`exp'3" 0 alcf2 0 alcf3 0 alcf4 1 alcf5 0 alcf6 0) /*
1369               /* at11("`exp'2" 1 "`exp'3" 0 alcf2 0 alcf3 0 alcf4 0 alcf5 1 alcf6 0) /*
1370               /* at12("`exp'2" 1 "`exp'3" 0 alcf2 0 alcf3 0 alcf4 0 alcf5 0 alcf6 1) /*
1371               /* at13("`exp'2" 0 "`exp'3" 1 alcf2 0 alcf3 0 alcf4 0 alcf5 0 alcf6 0) /*
1372               /* at14("`exp'2" 0 "`exp'3" 1 alcf2 1 alcf3 0 alcf4 0 alcf5 0 alcf6 0) /*
1373               /* at15("`exp'2" 0 "`exp'3" 1 alcf2 0 alcf3 1 alcf4 0 alcf5 0 alcf6 0) /*
1374               /* at16("`exp'2" 0 "`exp'3" 1 alcf2 0 alcf3 0 alcf4 1 alcf5 0 alcf6 0) /*
1375               /* at17("`exp'2" 0 "`exp'3" 1 alcf2 0 alcf3 0 alcf4 0 alcf5 1 alcf6 0) /*
1376               /* at18("`exp'2" 0 "`exp'3" 1 alcf2 0 alcf3 0 alcf4 0 alcf5 0 alcf6 1) /*
1377               timevar(timevar) ci userfunction(d`m') userfunctionvar(uf)
1378               local j = `m' + 12
1379               local dif = uf[1]
1380               local lb_dif = uf_lci[1]
1381               local ub_dif = uf_uci[1]
1382               preserve
1383               clear
1384               set obs 1
1385               gen sex = "`sex'"
1386               gen exp = "`exp'"
1387               gen dif = `dif'
1388               gen lb_dif = `lb_dif'
1389               gen ub_dif = `ub_dif'
1390               gen contr = "ct_`j'`m'"
1391               gen mod = "Alcohol frequency"
1392               append using `difalc', force
1393               save `difalc', replace
1394               restore
1395           }
1396           foreach t of numlist 11/16 21/26 31/36 {
1397               local s`t' = s`t'[1]
1398               local lb`t' = s`t'_lci[1]

```

```

1399         local ub` t' = s` t' _uci[1]
1400     }
1401     post stats    ("`sex'"') ("`out'"') ("`exp'"') (`npart') (`events') /*
1402     */           (`s11') (`lb11') (`ub11') (`s12') (`lb12') (`ub12') (`s13') (`lb13') (`ub13') (`s14') (`lb14') (`ub14')
1403     (`s15') (`lb15') (`ub15') (`s16') (`lb16') (`ub16') /*
1404     */           (`s21') (`lb21') (`ub21') (`s22') (`lb22') (`ub22') (`s23') (`lb23') (`ub23') (`s24') (`lb24') (`ub24')
1405     (`s25') (`lb25') (`ub25') (`s26') (`lb26') (`ub26') /*
1406     */           (`s31') (`lb31') (`ub31') (`s32') (`lb32') (`ub32') (`s33') (`lb33') (`ub33') (`s34') (`lb34') (`ub34')
1407     (`s35') (`lb35') (`ub35') (`s36') (`lb36') (`ub36')
1408 }
1409 }
1410 postclose stats
1411 preserve
1412 use `surv_alc', clear
1413 gen i = _n
1414 reshape long s lb ub, i(i) j(group)
1415 drop i
1416 gen level_exp      = "First"          if inlist(group, 11, 12, 13, 14, 15, 16)
1417 replace level_exp  = "Second"         if inlist(group, 21, 22, 23, 24, 25, 26)
1418 replace level_exp  = "Third"          if inlist(group, 31, 32, 33, 34, 35, 36)
1419 gen level_mod      = "Daily or almost daily" if inlist(group, 11, 21, 31)
1420 replace level_mod  = "Three or four times a week" if inlist(group, 12, 22, 32)
1421 replace level_mod  = "Once or twice a week"      if inlist(group, 13, 23, 33)
1422 replace level_mod  = "One to three times a month" if inlist(group, 14, 24, 34)
1423 replace level_mod  = "Special occasions only"    if inlist(group, 15, 25, 35)
1424 replace level_mod  = "Never"                  if inlist(group, 16, 26, 36)
1425 gen mod = "Alcohol frequency"
1426 order sex out exp level e mod level m s lb ub np ev group
1427 save "Results_part1\surv_alc", replace
1428
1429 use `difalc', clear
1430 gen level_mod      = "Daily or almost daily" if contr == "ct_131"
1431 replace level_mod  = "Three or four times a week" if contr == "ct_142"
1432 replace level_mod  = "Once or twice a week"      if contr == "ct_153"
1433 replace level_mod  = "One to three times a month" if contr == "ct_164"
1434 replace level_mod  = "Special occasions only"    if contr == "ct_175"
1435 replace level_mod  = "Never"                    if contr == "ct_186"
1436 drop contr
1437 save "Results_part1\survdif_alc", replace
1438 display "$S_TIME  $S_DATE"
1439 restore
1440
1441 #####
1442 *SENSITIVITY ANALYSIS: ROBUSTNESS TO COVARIATES INCLUSION*
1443
1444 *prepare database survival: survival difference, walking pace, men

```

```

1445   cls
1446   cd "Analysis"
1447   use "db3", clear
1448   xtilew gs3 = gs, within(sex) nq(3)
1449   drop bmi totw totm totv plew_t pleexerc_t rmeat fvscore cereal tv
1450   rename gs3 gt
1451   mdesc gt wp age0 sex ethn tws smok ndrugs hscore lnbmi
1452   egen float miss = rowmiss(gt wp age0 sex ethn tws ndrugs hscore lnbmi)
1453   tab miss, m
1454   drop if miss !=0
1455   drop miss
1456
1457   foreach var of varlist wp smok {
1458       tab `var', gen(`var')
1459   }
1460   drop wp wp1 smok smok1                                /*to facilitate wildcard stpm2 below in the loop*/
1461   gen timevar = 10 in 1
1462   stset time, id(n_eid) failure(all_death==1)
1463   tabstat all_death, statistics(count sum) by(sex)      /*participants and events*/
1464
1465   ****progressive adjustment****
1466   local adj0 = "wp2 wp3"
1467   local adj1 = "wp2 wp3 age0"
1468   local adj2 = "wp2 wp3 age0 ethn"
1469   local adj3 = "wp2 wp3 age0 ethn tws"
1470   local adj4 = "wp2 wp3 age0 ethn tws smok2 smok3"
1471   local adj5 = "wp2 wp3 age0 ethn tws smok2 smok3 ndrugs"
1472   local adj6 = "wp2 wp3 age0 ethn tws smok2 smok3 ndrugs hscore"
1473   local adj7 = "wp2 wp3 age0 ethn tws smok2 smok3 ndrugs hscore lnbmi"
1474
1475   preserve
1476   clear
1477   save "Results_part1\pro_haz", emptyok replace
1478   clear
1479   save "Results_part1\pro_sur", emptyok replace
1480   restore
1481
1482   forvalues k = 0/7 {
1483       stpm2 `adj`k'' if sex == 1, scale(hazard) df(4)
1484       preserve
1485       parmest, fast
1486       keep if (parm == "wp2" | parm == "wp3")
1487       gen adjl = `k'
1488       gen adjn = "`adj`k'"
1489       append using "Results_part1\pro_haz"
1490       save "Results_part1\pro_haz", replace
1491       restore
1492       preserve
1493       stpm2_standsurv, atvars(s1 s2 s3) /*
1494       */ at1(wp2 0 wp3 0) /*
1495       */ at2(wp2 1 wp3 0) /*

```

```

1496      */          at3(wp2 0 wp3 1) /*
1497      */ timevar(timevar) ci contrast(difference)
1498      keep in 1
1499      keep s1-_contrast3_1_uci
1500      gen adjl = `k'
1501      gen adjn = "`adj`k'"
1502      append using "Results_part1\pro_sur"
1503      save "Results_part1\pro_sur", replace
1504      restore
1505  }
1506
1507
1508  #####
1509  #####
1510  *P-values for main exposure and effect modifier*
1511  cls
1512  display "$S_TIME  $S_DATE"
1513  cd "Analysis"
1514  use "db3", clear
1515  xtilew gs3 = gs, within(sex) nq(3)
1516  gen grip = "sex" + string(sex) + " " + "gs" + string(gs3)
1517  sencode grip, gsort(grip) replace
1518  gen pace = "sex" + string(sex) + "_" + "pace" + string(wp)
1519  sencode pace, gsort(pace) replace
1520  describe, short
1521  drop bmi totw totm totv plew_t pleexerc_t rmeat fvscore cereal tv
1522  stset time, id(n eid) failure(all death==1)
1523
1524  ****categorical
1525  preserve
1526  clear
1527  tempfile models
1528  gen sex = .
1529  gen exp = .
1530  gen mod = .
1531  gen type = .
1532  gen p_exp = .
1533  gen p_mod = .
1534  save `models', emptyok replace
1535  restore
1536
1537  cls
1538  forvalues k = 0/1 {
1539      foreach exp in gs3 wp {
1540          foreach var of varlist smok alcf totw3 totm3 totv3 plew_t3 pleexerc_t3 /*cat*/ {
1541              stpm2 age0 ethn tws i.smok ndrugs i.`exp' i.`var' if sex == `k', scale(hazard) df(4) eform lininit /*ignored by
1542          stata if repeated exposure smok*/
1543              testparm i.`exp'
1544              local p_exp = r(p)

```

```

1544         testparm i.`var'
1545         local p_mod = r(p)
1546         preserve
1547         clear
1548         set obs 1
1549         gen sex = "`k'"
1550         gen exp = "`exp'"
1551         gen mod = "`var'"
1552         gen type = "cat"
1553         gen p_exp = `p_exp'
1554         gen p_mod = `p_mod'
1555         append using `models', force
1556         save `models', replace
1557         restore
1558     }
1559 }
1560 }
1561
1562
1563 ****continous
1564 cls
1565 foreach var of varlist wmeat lnrmear pmeat lnfv lncr lntv cpu sleep lnbmi hscore {
1566     rcsgen `var', df(3) gen(spl `var') orthog
1567 }
1568
1569 *no bmi (used and then removed as it is only in men interaction)
1570 forvalues k = 0/1 {
1571     foreach exp in gs3 wp {
1572         foreach var of varlist wmeat lnrmear pmeat lnfv lncr lntv cpu sleep lnbmi hscore /*cont*/ {
1573             stpm2 age0 ethn tws i.smok ndrugs i.`exp' c.spl_`var'* if sex == `k', scale(hazard) df(4) eform lininit
1574             testparm i.`exp'
1575             local p_exp = r(p)
1576             testparm c.spl_`var'*
1577             local p_mod = r(p)
1578             preserve
1579             clear
1580             set obs 1
1581             gen sex = "`k'"
1582             gen exp = "`exp'"
1583             gen mod = "`var'"
1584             gen type = "cont_spline"
1585             gen p_exp = `p_exp'
1586             gen p_mod = `p_mod'
1587             append using `models', force
1588             save `models', replace
1589         restore
1590     }
1591 }
1592 }
1593
1594 *bmi*walking_pace in men

```



```

1595 stpm2 age0 ethn tws i.smok ndrugs i.wp##c.spl_lnbmi* if sex == 1, scale(hazard) df(4) eform lininit
1596 testparm i.wp
1597 local p_exp = r(p)
1598 testparm c.spl_lnbmi*
1599 local p_mod = r(p)
1600 preserve
1601 clear
1602 set obs 1
1603 gen sex = "1"
1604 gen exp = "wp"
1605 gen mod = "lnbmi"
1606 gen type = "cont_spline_inter"
1607 gen p_exp = `p_exp'
1608 gen p_mod = `p_mod'
1609 gen i_bmi = 1
1610 append using `models', force
1611 save `models', replace
1612 restore
1613
1614 *combine results*
1615 use `models', clear
1616 drop if (mod == "lnbmi" & exp == "wp" & sex == "1" & i_bmi == .) /*from analysis without interaction*/
1617 replace sex = "Women" if sex == "0"
1618 replace sex = "Men" if sex == "1"
1619 foreach var of varlist sex-mod {
1620     sencode `var', replace
1621 }
1622 bys sex: tab exp mod
1623 drop i bmi
1624 foreach var of varlist sex-mod {
1625     sdecode `var', replace
1626 }
1627 save "Results_part1\p_values", replace
1628
1629 import excel "interactions.xls", sheet("Sheet1") firstrow clear
1630 drop type
1631 rename cutoff cutoff_int
1632 merge 1:1 sex exp mod using "Results_part1\p_values"
1633 drop _merge
1634 order out sex exp mod type aic aic_i dif2 p_int cutoff_int sig dif
1635 foreach var of varlist p_* {
1636     tostring `var', format(%7.3f) force replace
1637     replace `var' = "<0.001" if `var' == "0.000"
1638 }
1639 drop dif
1640 rename dif2 dif
1641 rename sig sig_int
1642 export excel using "Results_part1\interactions_pvalues.xls", firstrow(variables) replace
1643 display "$S_TIME $S_DATE"
1644
1645

```

```

1646 #####
1647 #####
1648 **Interaction WP*GS analysis*
1649 cls
1650 graph close _all
1651 graph drop _all
1652 cd "Analysis"
1653 use "db3", clear
1654 drop bmi totw totm totv plew t pleexerc t rmeat fvscore cereal tv
1655 describe, short
1656
1657 *Interaction: LRtest
1658 stset time, id(n_eid) failure(all_death==1)
1659 forvalues k = 0/1 {
1660     stpm2 age0 ethn tws i.smok ndrugs i.wp##c.gs if sex == `k', scale(hazard) df(4) eform lininit
1661     estimate store inter_`k'
1662     stpm2 age0 ethn tws i.smok ndrugs i.wp gs if sex == `k', scale(hazard) df(4) eform lininit
1663     estimate store nointer_`k'
1664 }
1665 lrtest (nointer_0) (inter_0)
1666 lrtest (nointer_1) (inter_1)
1667
1668 *Estimates
1669 foreach var of varlist wp smok {
1670     tab `var', gen(`var')
1671 }
1672
1673 tabstat gs, statistics(n p50 p5 p95) by(sex)
1674 histogram gs
1675 graph close _all
1676 local lev_gs 13(1.025)54 /*40 points estimation*/
1677 gen timevar = 10 in 1
1678
1679 tempfile survs_wpgs
1680 cap postclose stats
1681 postfile stats str5 sex str10 out str5 exp str5 mod npart events s1 lb1 ub1 s2 lb2 ub2 s3 lb3 ub3 d3 lbd3 ubd3 xval using
1682 `survs_wpgs'
1683
1683 forvalues sex = 0/1 {
1684     foreach mod of varlist gs {
1685         cap drop `mod's*
1686         rcsgen `mod', gen(`mod's) orthog df(3)
1687         global K`mod' `r(knots)'
1688         matrix M`mod' = r(R)
1689         foreach out in all_death {
1690             stset time, id(n_eid) failure(`out'==1)
1691             tabstat `out' if (`mod' !=. & sex == `sex'), statistics(count sum) save
1692             qui tabstatmat mx
1693             local npart = mx[1,1]

```

```

1694         local events = mx[2,1]
1695         foreach exp in wp {
1696             cap nois stpm2 age0 ethn tws smok2 smok3 ndrugs `exp'2 `exp'3 `mod's1 `mod's2 `mod's3 if sex == `sex', scale(hazard
) df(4) eform lininit
1697             forvalues k = `lev_`mod'' {
1698                 cap drop surv1* surv2* surv3* d2* d3*
1699                 rcsgen, scalar(`k') knots(`${K`mod}`) rmatrix(M`mod') gen(v)
1700                 stpm2_standsurv, atvars(surv1 surv2 surv3) /*
1701                 */ at1("`exp'2" 0 "`exp'3" 0 "`mod's1" `=v1' "`mod's2" `=v2' "`mod's3" `=v3') /*
1702                 */ at2("`exp'2" 1 "`exp'3" 0 "`mod's1" `=v1' "`mod's2" `=v2' "`mod's3" `=v3') /*
1703                 */ at3("`exp'2" 0 "`exp'3" 1 "`mod's1" `=v1' "`mod's2" `=v2' "`mod's3" `=v3') /*
1704                 */ timevar(timevar) ci contrast(difference) contrastvars(d2 d3)
1705                 local s1 = surv1[1]
1706                 local s2 = surv2[1]
1707                 local s3 = surv3[1]
1708                 local lb1 = surv1_lci[1]
1709                 local lb2 = surv2_lci[1]
1710                 local lb3 = surv3_lci[1]
1711                 local ub1 = surv1_uci[1]
1712                 local ub2 = surv2_uci[1]
1713                 local ub3 = surv3_uci[1]
1714                 local d3 = d3[1]
1715                 local lbd3 = d3_lci[1]
1716                 local ubd3 = d3_uci[1]
1717                 local xval = `k'
1718                 post stats ("`sex'") ("`out'") ("`exp'") ("`mod'") (`npart') (`events') (`s1') (`lb1') (`ub1') (`s2') (`lb2') (
`ub2') (`s3') (`lb3') (`ub3') (`d3') (`lbd3') (`ubd3') (`xval')
1719             }
1720         }
1721     }
1722 }
1723 }
1724 postclose stats
1725 use `survs_wpgs', clear
1726 replace sex = "Women" if sex == "0"
1727 replace sex = "Men" if sex == "1"
1728 drop if sex == "Women" & (xval>33.5 | xval<13) /*according to sex-specific 5th and 95th percentile of grip*/
1729 drop if sex == "Men" & (xval>54 | xval<25) /*according to sex-specific 5th and 95th percentile of grip*/
1730 foreach var of varlist d3 lbd3 ubd3 {
1731     replace `var' = `var'*1000/10 /*per 1000 persons per 10 years - to 1000 persons per year */
1732 }
1733

```