```
1  ********************************************************************************************************************************
   ********************************************************************
2  ***SYNTAX FOR "Socio-economic inequalities and trajectories of a novel multidimensional metric of
   Active and Healthy Ageing: the English Longitudinal Study of Ageing"***
3  ********************************************************************************************************************************
   ********************************************************************
4
5  * STATA version: 17.0, BE-Basic Edition
6
7  * STATA citation: StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX:
   StataCorp LLC.
8
9  * Data citation: Banks, J., Batty, G. David, Breedvelt, J., Coughlin, K., Crawford, R., Marmot, M.,
   Nazroo, J., Oldfield, Z., Steel, N., Steptoe, A., Wood, M., Zaninotto, P. (2021). English
   Longitudinal Study of Ageing: Waves 0-9, 1998-2019. [data collection]. 37th Edition. UK Data
   Service. SN: 5050, DOI: 10.5255/UKDA-SN-5050-24
10
11 * Data access statement: ELSA data from all waves are available through the UK Data Service
   (https://ukdataservice.ac.uk/). The main ELSA dataset is safeguarded and can be accessed via
   https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5050#!/access-data. More information
   on how to access ELSA, including the conditions of use, can be found on the UK Data Service website
   (https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5050#!/details) and the ELSA
   website (https://www.elsa-project.ac.uk/accessing-elsa-data).
12
13 * Date of data access/download (dd/mm/yyyy): 25/05/2022
14
15 * Project ID: 212810
16
17 * Data documentation: Documentation pertaining to ELSA (e.g., data dictionaries, questionnaires,
   technical reports, user guides) is available on the UK Data Service website
   (https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5050#!/documentation) and the ELSA
   website (https://www.elsa-project.ac.uk/data-and-documentation).
18
19 **********************
20 ***DATA PROCESSING***
21 **********************
22
23 * Change working directory - add pathname in between quotation marks for Windows
24 cd ""
25
26 * Variables Wave 2
27 use idauniq Heill Helim MmAlone MmHSS Hehelf w2nssec8 MmTrya MMWlkA MmTryb MMWlkB heada01 heada02
   heada03 heada04 heada05 heada06 heada07 heada08 heada09 heada10 headb01 headb02 headb03 headb04
   headb05 headb06 headb07 headb08 headb09 headb10 headb11 headb12 headb13 scorg01 scorg02 scorg03
   scorg04 scorg05 scorg06 scorg07 scorg08 scacta scactb scactc scactd scpt04 scpt05 sampsta dhager
   diagr indager DhSex DiSex indsex HeSmk HESka HeSkb HeSkc HeSkd HeSke HeSkf HeActa HeActb HeActc scako
    fqethnr scqola scqolb scqolc scqold scqole scqolf scqolg scqolh scqoli scqolj scqolk scqoll scqolm
   scqoln scqolo scqolp scqolq scqolr scqols PScedA PScedB PScedC PScedD PScedE PScedF PScedG PScedH
   scfeela scfeelb scfeelc CfDScr CfTest CfLisEn CfAni CfLisD MmWill MmSaf MmAvsp MmWala MmAid mmaidc
   CfDatD CfDatM CfDatY CfDay CfWrds MmSchs bhesmk bheska finstat w2wgt using wave_2_core_data_v4.dta
28 * Describe dataset
29 describe
30 * Sort from lowest to highest participant identifier (ID)
31 sort idauniq
32 * Rename variables to shorter or more convenient forms
33 rename MmSchs mmschs
34 rename MmWill mmwill
35 rename MmSaf mmsaf
36 rename MmAvsp mmavsp
37 rename MmWala mmwala
38 rename MmAid mmaid
39 rename CfDatD cfdatd
40 rename CfDatM cfdatm
```

```
41  rename CfDatY cfdaty
42  rename CfDay cfday
43  rename CfTest cftest
44  rename CfWrds cfwrds
45  rename Heill heill
46  rename Helim helim
47  rename MmAlone mmalone
48  rename MmHSS mmhss
49  rename Hehelf hehelf
50  rename MmTrya mmtrya
51  rename MMWlkA mmwlka
52  rename MmTryb mmtryb
53  rename MMWlkB mmwlkb
54  rename sampsta samptyp2
55  rename DhSex dhsex
56  rename DiSex disex
57  rename HeSmk hesmk
58  rename HESka heska
59  rename HeSkb heskb
60  rename HeSkc heskc
61  rename HeSkd heskd
62  rename HeSke heske
63  rename HeSkf heskf
64  rename HeActa heacta
65  rename HeActb heactb
66  rename HeActc heactc
67  rename PScedA psceda
68  rename PScedB pscedb
69  rename PScedC pscedc
70  rename PScedD pscedd
71  rename PScedE pscede
72  rename PScedF pscedf
73  rename PScedG pscedg
74  rename PScedH pscedh
75  rename CfDScr cfdscr
76  rename CfLisEn cflisen
77  rename CfAni cfani
78  rename CfLisD cflisd
79  rename w2nssec8 nssec8
80  rename finstat finstatw2
81  * Generate a new variable called wave and assign the number 2 to each observation (to designate Wave
    2)
82  gen wave = 2
83  * Save Wave 2 core dataset
84  save wave2.dta
85
86  * Variables Wave 3
87  use idauniq heill helim w3nssec8 mmalone mmhss hegenh mmtrya mmwlka mmtryb mmwlkb hemobwa hemobsi
    hemobch hemobcs hemobcl hemobst hemobre hemobpu hemobli hemobpi hemob96 headldr headlwa headlba
    headlea headlbe headlwc headlma headlpr headlsh headlph headlme headlho headlmo headl96 scorg01
    scorg02 scorg03 scorg04 scorg05 scorg06 scorg07 scorg08 scacta scactb scactc scactd scpt04 scpt05
    sampsta dhager diagr indager dhsex disex indsex hesmk heska heskb heskc heskd heske heskf chesmk
    cheska heacta heactb heactc scako fqethnr w3edqual scqola scqolb scqolc scqold scqole scqolf scqolg
    scqolh scqoli scqolj scqolk scqoll scqolm scqoln scqolo scqolp scqolq scqolr scqols psceda pscedb
    pscedc pscedd pscede pscedf pscedg pscedh scfeela scfeelb scfeelc cfdscr cftest cflisen cfani cflisd
    mmwill mmsaf mmavsp mmwala mmaid mmaidc cfdatd cfdatm cfdaty cfday cfwrds mmschs finstat w3xwgt using
     wave_3_elsa_data_v4.dta
88  * Describe dataset
89  describe
90  * Sort from lowest to highest participant ID
91  sort idauniq
92  * Rename variables to ensure consistency across waves
93  rename w3nssec8 nssec8
```

```
rename sampsta samptyp3
rename w3edqual edqual
rename finstat finstatw3
* Generate a new variable called wave and assign the number 3 to each observation (to designate Wave 3)
gen wave = 3
* Save Wave 3 core dataset
save wave3.dta

* Variables Wave 4
use idauniq heill helim w4nssec8 mmalone mmhss hehelf mmtrya mmwlka mmtryb mmwlkb hemobwa hemobsi hemobch hemobcs hemobcl hemobst hemobre hemobpu hemobli hemobpi hemob96 headldr headlwa headlba headlea headlbe headlwc headlma headlpr headlsh headlte headlme headlho headlmo headl96 scorg01 scorg02 scorg03 scorg04 scorg05 scorg06 scorg07 scorg08 scacta scactb scactc scactd scpt03 scpt04 samptyp diagr indager dhsex disex indsex hesmk heska heskb heskc heskd heske heskf dhesmk dheska heacta heactb heactc scako fqethnr w4edqual scqola scqolb scqolc scqold scqole scqolf scqolg scqolh scqoli scqolj scqolk scqoll scqolm scqoln scqolo scqolp scqolq scqolr scqols psceda pscedb pscedc pscedd pscede pscedf pscedg pscedh scfeela scfeelb scfeelc cfdscr cftest cflisen cfani cflisd mmwill mmsaf mmavsp mmwala mmaid mmaidc cfdatd cfdatm cfdaty cfday cfwrds mmschs finstat4 w4xwgt using wave_4_elsa_data_v3.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Rename variables to ensure consistency across waves
rename w4nssec8 nssec8
rename headlte headlph
rename scpt04 scpt05
rename scpt03 scpt04
rename w4edqual edqual
rename samptyp samptyp4
rename finstat4 finstatw4
* Generate a new variable called wave and assign the number 4 to each observation (to designate Wave 4)
gen wave = 4
* Save Wave 4 core dataset
save wave4.dta

* Variables Wave 5
use idauniq heill helim mmalone w5nssec8 mmhss hehelf mmtrya mmwlka mmtryb mmwlkb hemobwa hemobsi hemobch hemobcs hemobcl hemobst hemobre hemobpu hemobli hemobpi hemob96 headldr headlwa headlba headlea headlbe headlwc headlma headlpr headlsh headlte headlme headlho headlmo headl96 scorg01 scorg02 scorg03 scorg04 scorg05 scorg06 scorg07 scorg08 scacta scactb scactc scactd scpt03 scpt04 samptyp diagr indager dhsex disex indsex hesmk heska heskb heskc heskd heske heskf dhesmk dheska heacta heactb heactc scako fqethnr w5edqual scqola scqolb scqolc scqold scqole scqolf scqolg scqolh scqoli scqolj scqolk scqoll scqolm scqoln scqolo scqolp scqolq scqolr scqols psceda pscedb pscedc pscedd pscede pscedf pscedg pscedh scfeela scfeelb scfeelc cfdscr cftest cflisen cfani cflisd mmwill mmsaf mmavsp mmwala mmaid mmaidc cfdatd cfdatm cfdaty cfday cfwrds mmschs finstatw5 w5xwgt using wave_5_elsa_data_v4.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Rename variables to ensure consistency across waves
rename w5nssec8 nssec8
rename headlte headlph
rename scpt04 scpt05
rename scpt03 scpt04
rename w5edqual edqual
rename samptyp samptyp5
* Generate a new variable called wave and assign the number 5 to each observation (to designate Wave 5)
gen wave = 5
```

```
* Save Wave 5 core dataset
save wave5.dta

* Variables Wave 6
use idauniq Heill Helim w6nssec8 MmAlone MmHSS Hehelf MmTrya MMWlkA MmTryb MMWlkB hemobwa hemobsi
hemobch hemobcs hemobcl hemobst hemobre hemobpu hemobli hemobpi hemob96 headldr headlwa headlba
headlea headlbe headlwc headlma headlpr headlsh headlph headlme headlho headlmo headl96 scorg01
scorg02 scorg03 scorg04 scorg05 scorg06 scorg07 scorg08 scacta scactb scactc scactd scptr3 scptr4
samptyp indager DhSex DiSex indsex HeSmk HESka HeSkb HeSkc HeSkd HeSke HeSkf HeActa HeActb HeActc
scako Fqethnr scqola scqolb scqolc scqold scqole scqolf scqolg scqolh scqoli scqolj scqolk scqoll
scqolm scqoln scqolo scqolp scqolq scqolr scqols PScedA PScedB PScedC PScedD PScedE PScedF PScedG
PScedH scfeela scfeelb scfeelc CfDScr CfTest CfLisEn CfLisD MmWill MmSaf MmAvsp MmWala MmAid CfDatD
CfDatM CfDatY CfDay CfWrds MmSchs finstatw6 w6xwgt using wave_6_elsa_data_v2.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Rename variables to ensure consistency across waves
rename MmSchs mmschs
rename MmWill mmwill
rename MmSaf mmsaf
rename MmAvsp mmavsp
rename MmWala mmwala
rename MmAid mmaid
rename CfDatD cfdatd
rename CfDatM cfdatm
rename CfDatY cfdaty
rename CfDay cfday
rename CfTest cftest
rename CfWrds cfwrds
rename w6nssec8 nssec8
rename Heill heill
rename Helim helim
rename MmAlone mmalone
rename MmHSS mmhss
rename Hehelf hehelf
rename MmTrya mmtrya
rename MMWlkA mmwlka
rename MmTryb mmtryb
rename MMWlkB mmwlkb
rename scptr3 scpt04
rename scptr4 scpt05
rename DhSex dhsex
rename DiSex disex
rename HeSmk hesmk
rename HESka heska
rename HeSkb heskb
rename HeSkc heskc
rename HeSkd heskd
rename HeSke heske
rename HeSkf heskf
rename HeActa heacta
rename HeActb heactb
rename HeActc heactc
rename Fqethnr fqethnr
rename samptyp samptyp6
rename PScedA psceda
rename PScedB pscedb
rename PScedC pscedc
rename PScedD pscedd
rename PScedE pscede
rename PScedF pscedf
rename PScedG pscedg
```

```
rename PScedH pscedh
rename CfDScr cfdscr
rename CfLisEn cflisen
rename CfLisD cflisd
* Generate a new variable called wave and assign the number 6 to each observation (to designate Wave 6)
gen wave = 6
* Save Wave 6 core dataset
save wave6.dta

* Variables Wave 7
use idauniq Heill Helim NSSEC MmAlone MmHSS Hehelf MmTrya MMWlkA MmTryb MMWlkB hemobwa hemobsi hemobch hemobcs hemobcl hemobst hemobre hemobpu hemobli hemobpi hemob96 headldr headlwa headlba headlea headlbe headlwc headlma headlpr headlsh headlph headlme headlho headlmo headl96 scorg01 scorg02 scorg03 scorg04 scorg05 scorg06 scorg07 scorg08 scacta scactb scactc scactd scptr3 scptr4 samptyp indager DhSex DiSex indsex HeSmk HESka HeSkb HeSkc HeSkd HeSke HeSkf HeActa HeActb HeActc scako Fqethnr scqola scqolb scqolc scqold scqole scqolf scqolg scqolh scqoli scqolj scqolk scqoll scqolm scqoln scqolo scqolp scqolq scqolr scqols PScedA PScedB PScedC PScedD PScedE PScedF PScedG PScedH scfeela scfeelb scfeelc CfDScr CfTest CfLisEn CfAni CfLisD MmWill MmSaf MmAvsp MmWala MmAid CfDatD CfDatM CfDatY CfDay CfWrds MmSchs finstatw7 w7xwgt using wave_7_elsa_data.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Rename variables to ensure consistency across waves
rename MmSchs mmschs
rename MmWill mmwill
rename MmSaf mmsaf
rename MmAvsp mmavsp
rename MmWala mmwala
rename MmAid mmaid
rename CfDatD cfdatd
rename CfDatM cfdatm
rename CfDatY cfdaty
rename CfDay cfday
rename CfTest cftest
rename CfWrds cfwrds
rename Heill heill
rename Helim helim
rename MmAlone mmalone
rename MmHSS mmhss
rename Hehelf hehelf
rename MmTrya mmtrya
rename MMWlkA mmwlka
rename MmTryb mmtryb
rename MMWlkB mmwlkb
rename scptr3 scpt04
rename scptr4 scpt05
rename DhSex dhsex
rename DiSex disex
rename HeSmk hesmk
rename HESka heska
rename HeSkb heskb
rename HeSkc heskc
rename HeSkd heskd
rename HeSke heske
rename HeSkf heskf
rename HeActa heacta
rename HeActb heactb
rename HeActc heactc
rename Fqethnr fqethnr
rename samptyp samptyp7
rename PScedA psceda
```

```
245  rename PScedB pscedb
246  rename PScedC pscedc
247  rename PScedD pscedd
248  rename PScedE pscede
249  rename PScedF pscedf
250  rename PScedG pscedg
251  rename PScedH pscedh
252  rename CfDScr cfdscr
253  rename CfLisEn cflisen
254  rename CfAni cfani
255  rename CfLisD cflisd
256  * Generate a new variable called wave and assign the number 7 to each observation (to designate Wave
     7)
257  gen wave = 7
258  * Save Wave 7 core dataset
259  save wave7.dta
260
261  * Variables Wave 8
262  use idauniq heill helim w8nssec8 mmalone mmhss hehelf mmtrya mmwlka mmtryb mmwlkb hemobwa hemobsi
     hemobch hemobcs hemobcl hemobst hemobre hemobpu hemobli hemobpi hemob96 headldr headlwa headlba
     headlea headlbe headlwc headlma headlpr headlsh headlph headlme headlho headlmo headl96 scorgpo
     scorgnw scorgrl scorgch scorged scorgsc scorgsp scorg95 scacta scactb scactc scactd scptruk scptrab
     samptyp indager indsex hesmk heska heskb heskc heskd heske heskf heacta heactb heactc scako fqethnmr
     scqola scqolb scqolc scqold scqole scqolf scqolg scqolh scqoli scqolj scqolk scqoll scqolm scqoln
     scqolo scqolp scqolq scqolr scqols psceda pscedb pscedc pscedd pscede pscedf pscedg pscedh scfeela
     scfeelb scfeelc cfdscr cftest cflisen cfani cflisd mmwill mmsaf mmavsp mmwala mmaid cfdatd cfdatm
     cfdaty cfday cfwrds mmschs finstat w8xwgt using wave_8_elsa_data_eul_v2.dta
263  * Describe dataset
264  describe
265  * Sort from lowest to highest participant ID
266  sort idauniq
267  * Rename variables to ensure consistency across waves
268  rename w8nssec8 nssec8
269  rename scorgpo scorg01
270  rename scorgnw scorg02
271  rename scorgrl scorg03
272  rename scorgch scorg04
273  rename scorged scorg05
274  rename scorgsc scorg06
275  rename scorgsp scorg07
276  rename scorg95 scorg08
277  rename scptruk scpt04
278  rename scptrab scpt05
279  rename fqethnmr fqethnr
280  rename samptyp samptyp8
281  rename finstat finstatw8
282  * Generate a new variable called wave and assign the number 8 to each observation (to designate Wave
     8)
283  gen wave = 8
284  * Save Wave 8 core dataset
285  save wave8.dta
286
287  * Variables Wave 9
288  use idauniq heill helim w9nssec8 mmalone mmhss hehelf mmtrya mmwlka mmtryb mmwlkb hemobwa hemobsi
     hemobch hemobcs hemobcl hemobst hemobre hemobpu hemobli hemobpi hemob96 headldr headlwa headlba
     headlea headlbe headlwc headlma headlpr headlsh headlph headlme headlho headlmo headl96 scorgpo
     scorgnw scorgrl scorgch scorged scorgsc scorgsp scorg95 scacta scactb scactc scactd scptruk scptrab
     samptyp indager indsex hesmk heska heskb heskc heskd heske heskf heacta heactb heactc scalcm fqethnmr
      scqola scqolb scqolc scqold scqole scqolf scqolg scqolh scqoli scqolj scqolk scqoll scqolm scqoln
     scqolo scqolp scqolq scqolr scqols psceda pscedb pscedc pscedd pscede pscedf pscedg pscedh scfeela
     scfeelb scfeelc cfdscr cftest cflisen cfani cflisd mmwill mmsaf mmavsp mmwala mmaid cfdatd cfdatm
     cfdaty cfday cfwrds mmschs finstat w9xwgt using wave_9_elsa_data_eul_v1.dta
289  * Describe dataset
```

```
describe
* Sort from lowest to highest participant ID
sort idauniq
* Rename variables to ensure consistency across waves
rename w9nssec8 nssec8
rename scorgpo scorg01
rename scorgnw scorg02
rename scorgrl scorg03
rename scorgch scorg04
rename scorged scorg05
rename scorgsc scorg06
rename scorgsp scorg07
rename scorg95 scorg08
rename scptruk scpt04
rename scptrab scpt05
rename scalcm scako
rename fqethnmr fqethnr
rename samptyp samptyp9
rename finstat finstatw9
* Generate a new variable called wave and assign the number 9 to each observation (to designate Wave 9)
gen wave = 9
* Save Wave 9 core dataset
save wave9.dta

* Variables Wave 2 Financial Derived
use idauniq totwq5_bu_s using wave_2_financial_derived_variables.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Save Wave 2 financial derived dataset
save wave2financial.dta

* Variables Wave 2 Nurse
use idauniq mmssre mmssti mmstre mmstti mmftre2 mmrrre mmftti mmrrfti mmssna mmstna mmftna mmcrre mmcrna mmrrna cfib hdl trig ldl hscrp hba1c mmgsd1 mmgsn1 mmgsd2 mmgsn2 mmgsd3 mmgsn3 mmgswil mmgsdom mmgssta mmgstp mmgsres mmbcsc mmsssc mmstsc mmftsc mmcrav mmcrsc mmrrsc mmrrtti mmrroc using wave_2_nurse_data_v2.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Save Wave 2 nurse dataset
save wave2nurse.dta

* Variables Wave 2 Derived
use idauniq w2edqual using wave_2_derived_variables.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Rename variables to shorter or more convenient forms
rename w2edqual edqual
* Save Wave 2 derived dataset
save wave2derived.dta

* Variables Wave 3 Financial Derived
use idauniq totwq5_bu_s using wave_3_financial_derived_variables.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
```

```
* Save Wave 3 financial derived dataset
save wave3financial.dta

* Variables Wave 4 Financial Derived
use idauniq totwq5_bu_s using wave_4_financial_derived_variables.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Save Wave 4 financial derived dataset
save wave4financial.dta

* Variables Wave 4 Nurse
use idauniq mmssre mmssti mmstre mmstti mmftre2 mmrrre mmftti mmrrfti mmssna mmstna mmftna mmcrre
mmcrna mmrrna cfib hdl trig ldl hscrp hba1c mmgsd1 mmgsn1 mmgsd2 mmgsn2 mmgsd3 mmgsn3 mmgswil mmgsdom
 mmgssta mmgstp mmgsres mmbcsc mmsssc mmstsc mmftsc mmcrav mmcrsc mmrrsc mmrrtti mmrroc using
wave_4_nurse_data.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Save Wave 4 nurse dataset
save wave4nurse.dta

* Variables Wave 5 Financial Derived
use idauniq totwq5_bu_s using wave_5_financial_derived_variables.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Save Wave 5 financial derived dataset
save wave5financial.dta

* Variables Wave 6 Financial Derived
use idauniq totwq5_bu_s using wave_6_financial_derived_variables.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Save Wave 6 financial derived dataset
save wave6financial.dta

* Variables Wave 6 Nurse
use idauniq mmssre mmssti mmstre mmstti MmFTRe MMFTRE2 mmrrre mmftti mmrrfti mmssna mmstna mmftna
mmcrre mmcrna mmrrna cfib hdl trig ldl hscrp hba1c mmgsd1 mmgsn1 mmgsd2 mmgsn2 mmgsd3 mmgsn3 mmgswil
mmgsdom mmgssta mmgstp mmgsres mmbcsc mmsssc mmstsc mmftsc mmcrav mmcrsc mmrrsc mmrrtti MMRROC using
wave_6_elsa_nurse_data_v2.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Rename variables to ensure consistency across waves
rename MMRROC mmrroc
rename MmFTRe mmftre
rename MMFTRE2 mmftre2
* Save Wave 6 nurse dataset
save wave6nurse.dta

* Variables Wave 6 Derived
use idauniq edqual using wave_6_ifs_derived_variables.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
```

```
406  sort idauniq
407  * Save Wave 6 derived dataset
408  save wave6derived.dta
409
410  * Variables Wave 7 Financial Derived
411  use idauniq totwq5_bu_s using wave_7_financial_derived_variables.dta
412  * Describe dataset
413  describe
414  * Sort from lowest to highest participant ID
415  sort idauniq
416  * Save Wave 7 financial derived dataset
417  save wave7financial.dta
418
419  * Variables Wave 7 Derived
420  use idauniq edqual using wave_7_ifs_derived_variables.dta
421  * Describe dataset
422  describe
423  * Sort from lowest to highest participant ID
424  sort idauniq
425  * Save Wave 7 derived dataset
426  save wave7derived.dta
427
428  * Variables Wave 8 Financial Derived
429  use idauniq totwq5_bu_s using wave_8_elsa_financial_dvs_eul_v1.dta
430  * Describe dataset
431  describe
432  * Sort from lowest to highest participant ID
433  sort idauniq
434  * Save Wave 8 financial derived dataset
435  save wave8financial.dta
436
437  * Variables Wave 8 Derived
438  use idauniq edqual using wave_8_elsa_ifs_dvs_eul_v1.dta
439  * Describe dataset
440  describe
441  * Sort from lowest to highest participant ID
442  sort idauniq
443  * Save Wave 8 derived dataset
444  save wave8derived.dta
445
446  * Variables Wave 8-9 Nurse
447  use idauniq wave cfib hdl trig ldl hscrp hba1c mmgsd1 mmgsn1 mmgsd2 mmgsn2 mmgsd3 mmgsn3 mmgswil
     mmgsdom mmgssta mmgstp mmgsres using elsa_nurse_w8w9_data_eul.dta
448  * Describe dataset
449  describe
450  * Sort from lowest to highest participant ID
451  sort idauniq
452  * Save Wave 8-9 nurse dataset
453  save wave89nurse.dta
454  * Keep data from Wave 8 only
455  keep if wave==8
456  * Save Wave 8 nurse dataset
457  save wave8nurse.dta
458  * Use Wave 8-9 nurse dataset
459  use wave89nurse.dta
460  * Keep data from Wave 9 only
461  keep if wave==9
462  * Save Wave 9 nurse dataset
463  save wave9nurse.dta
464
465  * Variables Wave 9 Financial Derived
466  use idauniq totwq5_bu_s using wave_9_financial_derived_variables.dta
467  * Describe dataset
```

```
describe
* Sort from lowest to highest participant ID
sort idauniq
* Save Wave 9 financial derived dataset
save wave9financial.dta

* Wave 9 Derived variables
use idauniq edqual using wave_9_ifs_derived_variables.dta
* Describe dataset
describe
* Sort from lowest to highest participant ID
sort idauniq
* Save Wave 9 derived dataset
save wave9derived.dta

* Wave 2 complete data
* Merge core, financial, nurse, and derived datasets for Wave 2 using the participant ID
* Use Wave 2 core dataset
use wave2.dta
* One-to-one merge of data in memory with wave2financial.dta on participant ID
merge 1:1 idauniq using wave2financial.dta, generate (merge_financial2)
* Overwrite Wave 2 dataset, by replacing the previously saved file
save wave2.dta, replace
* Use the newly saved file for Wave 2
use wave2.dta
* One-to-one merge of data in memory with wave2nurse.dta on participant ID
merge 1:1 idauniq using wave2nurse.dta, generate (merge_nurse2)
* Overwrite Wave 2 dataset, by replacing the previously saved file
save wave2.dta, replace
* Use the newly saved file for Wave 2
use wave2.dta
* One-to-one merge of data in memory with wave2derived.dta on participant ID
merge 1:1 idauniq using wave2derived.dta, generate (merge_derived2)
* Sort from lowest to highest participant ID
sort idauniq
* Overwrite Wave 2 dataset, by replacing the previously saved file
save wave2.dta, replace

* Wave 3 complete data
* Merge core and financial datasets for Wave 3 using the participant ID
* Use Wave 3 core dataset
use wave3.dta
* One-to-one merge of data in memory with wave3financial.dta on participant ID
merge 1:1 idauniq using wave3financial.dta, generate (merge_financial3)
* Sort from lowest to highest participant ID
sort idauniq
* Overwrite Wave 3 dataset, by replacing the previously saved file
save wave3.dta, replace

* Wave 4 complete data
* Merge core, financial, and nurse datasets for Wave 4 using the participant ID
* Use Wave 4 core dataset
use wave4.dta
* One-to-one merge of data in memory with wave4financial.dta on participant ID
merge 1:1 idauniq using wave4financial.dta, generate (merge_financial4)
* Overwrite Wave 4 dataset, by replacing the previously saved file
save wave4.dta, replace
* Use the newly saved file for Wave 4
use wave4.dta
* One-to-one merge of data in memory with wave4nurse.dta on participant ID
merge 1:1 idauniq using wave4nurse.dta, generate (merge_nurse4)
* Sort from lowest to highest participant ID
sort idauniq
```

```
* Overwrite Wave 4 dataset, by replacing the previously saved file
save wave4.dta, replace

* Wave 5 complete data
* Merge core and financial datasets for Wave 5 using the participant ID
* Use Wave 5 core dataset
use wave5.dta
* One-to-one merge of data in memory with wave5financial.dta on participant ID
merge 1:1 idauniq using wave5financial.dta, generate (merge_financial5)
* Sort from lowest to highest participant ID
sort idauniq
* Overwrite Wave 5 dataset, by replacing the previously saved file
save wave5.dta, replace

* Wave 6 complete data
* Merge core, financial, nurse, and derived datasets for Wave 6 using the participant ID
* Use Wave 6 core dataset
use wave6.dta
* One-to-one merge of data in memory with wave6financial.dta on participant ID
merge 1:1 idauniq using wave6financial.dta, generate (merge_financial6)
* Overwrite Wave 6 dataset, by replacing the previously saved file
save wave6.dta, replace
* Use the newly saved file for Wave 6
use wave6.dta
* One-to-one merge of data in memory with wave6nurse.dta on participant ID
merge 1:1 idauniq using wave6nurse.dta, generate (merge_nurse6)
* Overwrite Wave 6 dataset, by replacing the previously saved file
save wave6.dta, replace
* Use the newly saved file for Wave 6
use wave6.dta
* One-to-one merge of data in memory with wave6derived.dta on participant ID
merge 1:1 idauniq using wave6derived.dta, generate (merge_derived6)
* Sort from lowest to highest participant ID
sort idauniq
* Overwrite Wave 6 dataset, by replacing the previously saved file
save wave6.dta, replace

* Wave 7 complete data
* Merge core, financial, and derived datasets for Wave 7 using the participant ID
* Use Wave 7 core dataset
use wave7.dta
* One-to-one merge of data in memory with wave7financial.dta on participant ID
merge 1:1 idauniq using wave7financial.dta, generate (merge_financial7)
* Overwrite Wave 7 dataset, by replacing the previously saved file
save wave7.dta, replace
* Use the newly saved file for Wave 7
use wave7.dta
* One-to-one merge of data in memory with wave7derived.dta on participant ID
merge 1:1 idauniq using wave7derived.dta, generate (merge_derived7)
* Sort from lowest to highest participant ID
sort idauniq
* Overwrite Wave 7 dataset, by replacing the previously saved file
save wave7.dta, replace

* Wave 8 complete data
* Merge core, financial, nurse, and derived datasets for Wave 8 using the participant ID
* Use Wave 8 core dataset
use wave8.dta
* One-to-one merge of data in memory with wave8financial.dta on participant ID
merge 1:1 idauniq using wave8financial.dta, generate (merge_financial8)
* Overwrite Wave 8 dataset, by replacing the previously saved file
save wave8.dta, replace
* Use the newly saved file for Wave 8
```

```
use wave8.dta
* One-to-one merge of data in memory with wave8nurse.dta on participant ID
merge 1:1 idauniq using wave8nurse.dta, generate (merge_nurse8)
* Sort from lowest to highest participant ID
sort idauniq
* Overwrite Wave 8 dataset, by replacing the previously saved file
save wave8.dta, replace
* Use the newly saved file for Wave 8
use wave8.dta
* One-to-one merge of data in memory with wave8derived.dta on participant ID
merge 1:1 idauniq using wave8derived.dta, generate (merge_derived8)
* Sort from lowest to highest participant ID
sort idauniq
* Overwrite Wave 8 dataset, by replacing the previously saved file
save wave8.dta, replace

* Wave 9 complete data
* Merge core, financial, nurse, and derived datasets for Wave 9 using the participant ID
* Use Wave 9 core dataset
use wave9.dta
* One-to-one merge of data in memory with wave9financial.dta on participant ID
merge 1:1 idauniq using wave9financial.dta, generate (merge_financial9)
* Overwrite Wave 9 dataset, by replacing the previously saved file
save wave9.dta, replace
* Use the newly saved file for Wave 9
use wave9.dta
* One-to-one merge of data in memory with wave9derived.dta on participant ID
merge 1:1 idauniq using wave9derived.dta, generate (merge_derived9)
* Sort from lowest to highest participant ID
sort idauniq
* Overwrite Wave 9 dataset, by replacing the previously saved file
save wave9.dta, replace
* Use the newly saved file for Wave 9
use wave9.dta
* One-to-one merge of data in memory with wave9nurse.dta on participant ID
merge 1:1 idauniq using wave9nurse.dta, generate (merge_nurse9)
* Sort from lowest to highest participant ID
sort idauniq
* Overwrite Wave 9 dataset, by replacing the previously saved file
save wave9.dta, replace

* Append Wave 3 dataset to Wave 2 dataset
use wave2.dta
append using wave3.dta
* Sort by participant ID and wave (lowest to highest)
sort idauniq wave
* Append Wave 4 dataset
append using wave4.dta
* Sort by participant ID and wave (lowest to highest)
sort idauniq wave
* Append Wave 5 dataset
append using wave5.dta
* Sort by participant ID and wave (lowest to highest)
sort idauniq wave
* Append Wave 6 dataset
append using wave6.dta
* Sort by participant ID and wave (lowest to highest)
sort idauniq wave
* Append Wave 7 dataset
append using wave7.dta
* Sort by participant ID and wave (lowest to highest)
sort idauniq wave
* Append Wave 8 dataset
```

```
append using wave8.dta
* Sort by participant ID and wave (lowest to highest)
sort idauniq wave
* Append Wave 9 dataset
append using wave9.dta
* Sort by participant ID and wave (lowest to highest)
sort idauniq wave
* Assign a number in ascending order to each row of observations
gen ascnr = _n

* Organising dataset
* Generate a variable that assigns the observation number (i.e., 1 for first data collection timepoint, 2 for second data collection timepoint etc.) to each row by participant ID
bysort idauniq (wave): gen obsnr = _n
* Generate a variable that assigns the number of total observations to each row of data for a given participant
bysort idauniq: gen obscount = _N
* Check how many participants have data at 1 to 8 timepoints - the "if obsnr==1" statement is used to prevent participants with data at more than one timepoint from contributing to the counts more than once
tabulate obscount if obsnr==1
* Generate a variable that assigns the number 1 to the row representing participants' first observation
bysort idauniq (wave): gen first = 1 if _n==1
* Generate a variable that assigns the number 1 to the row representing participants' last observation
bysort idauniq (wave): gen last = 1 if _n==_N
* Generate a variable that assigns the number 1 to the row representing participants' first observation if this corresponds to Wave 2 (baseline)
bysort idauniq (wave): gen firstwave = 1 if obsnr==1 & wave==2
* Carry the value of this last variable forwards to the remainder of a participant's observations
bysort idauniq: gen firstwave_cons = firstwave[1]
* Install unique command
ssc install unique
* Count total number of participants and observations
unique idauniq
* 19,807 individuals, 80,750 observations
* Save dataset with a new name
save raw.dta

* Keep if participant is a core member (include core members who had a proxy or partial interview or those who had been interviewed in an institution)
keep if finstatw2=="C1CM" | inlist(finstatw3,"C1CM","C3CM") | inlist(finstatw4,"C1CM","C3CM","C4CM") | inlist(finstatw5,"C1CM","C3CM","C4CM") | inlist(finstatw6,1,7,14,25) | inlist(finstatw7,1,7,14,25,33) | inlist(finstatw8,1,7,14,25,33) | inlist(finstatw9,1,7,14,25,33,48)
* Count total number of participants and observations
unique idauniq
* 15,022 individuals, 68,496 observations
* Replace age = 90 if participant is aged 90+ years (collapsed in ELSA and coded as 99 at Wave 2, 3, and 4)
replace indager = 90 if indager==99 & inlist(wave,2,3,4)
* Replace age = 90 if participant is aged 90+ years (collapsed in ELSA and coded as -7 at Wave 5, 6, 7, 8, and 9)
replace indager = 90 if indager==-7 & inlist(wave,5,6,7,8,9)
* Save dataset with a new name
save data.dta

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* [b37-b40] Socio-cultural trips (reversed)
* [b7-b8] Holidays
* Replace variables as missing for any missing cases (coded as negative numbers in the ELSA dataset)
* [b37] How often respondent goes to the cinema
replace scacta = . if scacta<0
* [b38] How often respondent eats out of the house
```

```
replace scactb = . if scactb<0
* [b39] How often respondent goes to an art gallery or museum
replace scactc = . if scactc<0
* [b40] How often respondent goes to the theatre, a concert, or the opera
replace scactd = . if scactd<0
* [b7] Respondent has taken a holiday in the UK in the last 12 months
replace scpt04 = . if scpt04<0
* [b8] Respondent has taken a holiday abroad in the last 12 months
replace scpt05 = . if scpt05<0

* Reverse the negatively framed variables (this creates new variables and adds the "rev" prefix to the original variable names)
revrs scacta
revrs scactb
revrs scactc
revrs scactd
* Collapse categories with a small number of participants
replace revscacta = 5 if revscacta == 6
replace revscactc = 5 if revscactc == 6
replace revscactd = 5 if revscactd == 6
replace revscactb = 1 if revscactb ==2
replace revscactb = 2 if revscactb ==3
replace revscactb = 3 if revscactb ==4
replace revscactb = 4 if revscactb ==5
replace revscactb = 5 if revscactb ==6

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* [b26-b36] (+ 2 omitted) ADL and IADL disabilities (reversed)
* Replace variables as missing for any missing cases (coded as negative numbers in the ELSA dataset)
replace heada01 = . if heada01<0
replace heada02 = . if heada02<0
replace heada03 = . if heada03<0
replace heada04 = . if heada04<0
replace heada05 = . if heada05<0
replace heada06 = . if heada06<0
replace heada07 = . if heada07<0
replace heada08 = . if heada08<0
replace heada09 = . if heada09<0
replace heada10 = . if heada10<0
replace headb01 = . if headb01<0
replace headb02 = . if headb02<0
replace headb03 = . if headb03<0
replace headb04 = . if headb04<0
replace headb05 = . if headb05<0
replace headb06 = . if headb06<0
replace headb07 = . if headb07<0
replace headb08 = . if headb08<0
replace headb09 = . if headb09<0
replace headb10 = . if headb10<0
replace headb11 = . if headb11<0
replace headb12 = . if headb12<0
replace headb13 = . if headb13<0

replace headldr = . if headldr<0
replace headlwa = . if headlwa<0
replace headlba = . if headlba<0
replace headlea = . if headlea<0
replace headlbe = . if headlbe<0
replace headlwc = . if headlwc<0
replace headlma = . if headlma<0
replace headlpr = . if headlpr<0
replace headlsh = . if headlsh<0
replace headlph = . if headlph<0
```

```
771  replace headlme = . if headlme<0
772  replace headlho = . if headlho<0
773  replace headlmo = . if headlmo<0
774  replace headl96 = . if headl96<0
775
776  * ADL
777  * [b26] ADL: difficulty dressing, including putting on shoes and socks
778  replace headldr = 2 if headldr == 0
779  * Assign the number 0 if the participant reported difficulties performing the first listed activity
780  replace headldr = 0 if headldr == 1
781  * Assign the number 1 if the participant reported no difficulties performing the first listed activity
782  replace headldr = 1 if headldr == 2
783  * Assign the number 0 if the participant reported difficulties performing the first listed activity
784  replace headldr = 0 if (headb01 == 1 | headb02 == 1 | headb03 == 1 | headb04 == 1 | headb05 == 1 |
     headb06 == 1 | headb07 == 1 | headb08 == 1 | headb09 == 1 | headb10 == 1 | headb11 == 1 | headb12 ==
     1 | headb13 == 1) & wave==2
785  * Assign the number 1 if the participant reported any answer other than the listed activity in
     headb01-headb13 and data are not missing
786  replace headldr = 1 if inlist(headb01,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
787  replace headldr = 1 if inlist(headb02,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
788  replace headldr = 1 if inlist(headb03,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
789  replace headldr = 1 if inlist(headb04,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
790  replace headldr = 1 if inlist(headb05,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
791  replace headldr = 1 if inlist(headb06,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
792  replace headldr = 1 if inlist(headb07,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
793  replace headldr = 1 if inlist(headb08,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
794  replace headldr = 1 if inlist(headb09,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
795  replace headldr = 1 if inlist(headb10,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
796  replace headldr = 1 if inlist(headb11,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
797  replace headldr = 1 if inlist(headb12,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
798  replace headldr = 1 if inlist(headb13,2,3,4,5,6,7,8,9,10,11,12,13,96) & headldr!=0
799
800  * [b27] ADL: difficulty walking across a room
801  replace headlwa = 2 if headlwa == 0
802  * Assign the number 0 if the participant reported difficulties performing the second listed activity
803  replace headlwa = 0 if headlwa == 1
804  * Assign the number 1 if the participant reported no difficulties performing the second listed
     activity
805  replace headlwa = 1 if headlwa == 2
806  * Assign the number 0 if the participant reported difficulties performing the second listed activity
807  replace headlwa = 0 if (headb01 == 2 | headb02 == 2 | headb03 == 2 | headb04 == 2 | headb05 == 2 |
     headb06 == 2 | headb07 == 2 | headb08 == 2 | headb09 == 2 | headb10 == 2 | headb11 == 2 | headb12 ==
     2 | headb13 == 2) & wave==2
808  * Assign the number 1 if the participant reported no difficulties performing the second listed
     activity
809  replace headlwa = 1 if inlist(headb01,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
810  replace headlwa = 1 if inlist(headb02,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
811  replace headlwa = 1 if inlist(headb03,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
812  replace headlwa = 1 if inlist(headb04,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
813  replace headlwa = 1 if inlist(headb05,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
814  replace headlwa = 1 if inlist(headb06,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
815  replace headlwa = 1 if inlist(headb07,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
816  replace headlwa = 1 if inlist(headb08,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
817  replace headlwa = 1 if inlist(headb09,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
818  replace headlwa = 1 if inlist(headb10,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
819  replace headlwa = 1 if inlist(headb11,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
820  replace headlwa = 1 if inlist(headb12,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
821  replace headlwa = 1 if inlist(headb13,1,3,4,5,6,7,8,9,10,11,12,13,96) & headlwa!=0
822
823  * [b28] ADL: difficulty bathing or showering
824  replace headlba = 2 if headlba == 0
825  * Assign the number 0 if the participant reported difficulties performing the third listed activity
826  replace headlba = 0 if headlba == 1
```

```
827 * Assign the number 1 if the participant reported no difficulties performing the third listed activity
828 replace headlba = 1 if headlba == 2
829 * Assign the number 0 if the participant reported difficulties performing the third listed activity
830 replace headlba = 0 if (headb01 == 3 | headb02 == 3 | headb03 == 3 | headb04 == 3 | headb05 == 3 |
    headb06 == 3 | headb07 == 3 | headb08 == 3 | headb09 == 3 | headb10 == 3 | headb11 == 3 | headb12 ==
    3 | headb13 == 3) & wave==2
831 * Assign the number 1 if the participant reported no difficulties performing the third listed activity
832 replace headlba = 1 if inlist(headb01,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
833 replace headlba = 1 if inlist(headb02,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
834 replace headlba = 1 if inlist(headb03,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
835 replace headlba = 1 if inlist(headb04,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
836 replace headlba = 1 if inlist(headb05,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
837 replace headlba = 1 if inlist(headb06,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
838 replace headlba = 1 if inlist(headb07,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
839 replace headlba = 1 if inlist(headb08,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
840 replace headlba = 1 if inlist(headb09,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
841 replace headlba = 1 if inlist(headb10,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
842 replace headlba = 1 if inlist(headb11,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
843 replace headlba = 1 if inlist(headb12,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
844 replace headlba = 1 if inlist(headb13,1,2,4,5,6,7,8,9,10,11,12,13,96) & headlba!=0
845
846 * [b29] ADL: difficulty eating, such as cutting up your food
847 replace headlea = 2 if headlea == 0
848 * Assign the number 0 if the participant reported difficulties performing the fourth listed activity
849 replace headlea = 0 if headlea == 1
850 * Assign the number 1 if the participant reported no difficulties performing the fourth listed
    activity
851 replace headlea = 1 if headlea == 2
852 * Assign the number 0 if the participant reported difficulties performing the fourth listed activity
853 replace headlea = 0 if (headb01 == 4 | headb02 == 4 | headb03 == 4 | headb04 == 4 | headb05 == 4 |
    headb06 == 4 | headb07 == 4 | headb08 == 4 | headb09 == 4 | headb10 == 4 | headb11 == 4 | headb12 ==
    4 | headb13 == 4) & wave==2
854 * Assign the number 1 if the participant reported no difficulties performing the fourth listed
    activity
855 replace headlea = 1 if inlist(headb01,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
856 replace headlea = 1 if inlist(headb02,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
857 replace headlea = 1 if inlist(headb03,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
858 replace headlea = 1 if inlist(headb04,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
859 replace headlea = 1 if inlist(headb05,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
860 replace headlea = 1 if inlist(headb06,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
861 replace headlea = 1 if inlist(headb07,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
862 replace headlea = 1 if inlist(headb08,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
863 replace headlea = 1 if inlist(headb09,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
864 replace headlea = 1 if inlist(headb10,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
865 replace headlea = 1 if inlist(headb11,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
866 replace headlea = 1 if inlist(headb12,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
867 replace headlea = 1 if inlist(headb13,1,2,3,5,6,7,8,9,10,11,12,13,96) & headlea!=0
868
869 * [b30] ADL: difficulty getting in or out of bed
870 replace headlbe = 2 if headlbe == 0
871 * Assign the number 0 if the participant reported difficulties performing the fifth listed activity
872 replace headlbe = 0 if headlbe == 1
873 * Assign the number 1 if the participant reported no difficulties performing the fifth listed activity
874 replace headlbe = 1 if headlbe == 2
875 * Assign the number 0 if the participant reported difficulties performing the fifth listed activity
876 replace headlbe = 0 if (headb01 == 5 | headb02 == 5 | headb03 == 5 | headb04 == 5 | headb05 == 5 |
    headb06 == 5 | headb07 == 5 | headb08 == 5 | headb09 == 5 | headb10 == 5 | headb11 == 5 | headb12 ==
    5 | headb13 == 5) & wave==2
877 * Assign the number 1 if the participant reported no difficulties performing the fifth listed activity
878 replace headlbe = 1 if inlist(headb01,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
879 replace headlbe = 1 if inlist(headb02,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
880 replace headlbe = 1 if inlist(headb03,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
881 replace headlbe = 1 if inlist(headb04,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
```

```
replace headlbe = 1 if inlist(headb05,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
replace headlbe = 1 if inlist(headb06,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
replace headlbe = 1 if inlist(headb07,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
replace headlbe = 1 if inlist(headb08,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
replace headlbe = 1 if inlist(headb09,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
replace headlbe = 1 if inlist(headb10,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
replace headlbe = 1 if inlist(headb11,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
replace headlbe = 1 if inlist(headb12,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0
replace headlbe = 1 if inlist(headb13,1,2,3,4,6,7,8,9,10,11,12,13,96) & headlbe!=0

* [b31] ADL: difficulty using the toilet, including getting up or down
replace headlwc = 2 if headlwc == 0
* Assign the number 0 if the participant reported difficulties performing the sixth listed activity
replace headlwc = 0 if headlwc == 1
* Assign the number 1 if the participant reported no difficulties performing the sixth listed activity
replace headlwc = 1 if headlwc == 2
* Assign the number 0 if the participant reported difficulties performing the sixth listed activity
replace headlwc = 0 if (headb01 == 6 | headb02 == 6 | headb03 == 6 | headb04 == 6 | headb05 == 6 |
headb06 == 6 | headb07 == 6 | headb08 == 6 | headb09 == 6 | headb10 == 6 | headb11 == 6 | headb12 ==
6 | headb13 == 6) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the sixth listed activity
replace headlwc = 1 if inlist(headb01,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb02,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb03,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb04,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb05,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb06,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb07,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb08,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb09,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb10,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb11,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb12,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0
replace headlwc = 1 if inlist(headb13,1,2,3,4,5,7,8,9,10,11,12,13,96) & headlwc!=0

* IADL
* [b32] IADL: difficulty using a map to figure out how to get around in a strange place
replace headlma = 2 if headlma == 0
* Assign the number 0 if the participant reported difficulties performing the first listed activity
replace headlma = 0 if headlma == 1
* Assign the number 1 if the participant reported no difficulties performing the first listed activity
replace headlma = 1 if headlma == 2
* Assign the number 0 if the participant reported difficulties performing the first listed activity
replace headlma = 0 if (headb01 == 7 | headb02 == 7 | headb03 == 7 | headb04 == 7 | headb05 == 7 |
headb06 == 7 | headb07 == 7 | headb08 == 7 | headb09 == 7 | headb10 == 7 | headb11 == 7 | headb12 ==
7 | headb13 == 7) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the first listed activity
replace headlma = 1 if inlist(headb01,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb02,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb03,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb04,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb05,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb06,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb07,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb08,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb09,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb10,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb11,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb12,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0
replace headlma = 1 if inlist(headb13,1,2,3,4,5,6,8,9,10,11,12,13,96) & headlma!=0

* [b33] IADL: difficulty preparing a hot meal
replace headlpr = 2 if headlpr == 0
```

```
* Assign the number 0 if the participant reported difficulties performing the second listed activity
replace headlpr = 0 if headlpr == 1
* Assign the number 1 if the participant reported no difficulties performing the second listed activity
replace headlpr = 1 if headlpr == 2
* Assign the number 0 if the participant reported difficulties performing the second listed activity
replace headlpr = 0 if (headb01 == 8 | headb02 == 8 | headb03 == 8 | headb04 == 8 | headb05 == 8 | headb06 == 8 | headb07 == 8 | headb08 == 8 | headb09 == 8 | headb10 == 8 | headb11 == 8 | headb12 == 8 | headb13 == 8) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the second listed activity
replace headlpr = 1 if inlist(headb01,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb02,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb03,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb04,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb05,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb06,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb07,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb08,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb09,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb10,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb11,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb12,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0
replace headlpr = 1 if inlist(headb13,1,2,3,4,5,6,7,9,10,11,12,13,96) & headlpr!=0

* [b34] IADL: difficulty shopping for groceries
replace headlsh = 2 if headlsh == 0
* Assign the number 0 if the participant reported difficulties performing the third listed activity
replace headlsh = 0 if headlsh == 1
* Assign the number 1 if the participant reported no difficulties performing the third listed activity
replace headlsh = 1 if headlsh == 2
* Assign the number 0 if the participant reported difficulties performing the third listed activity
replace headlsh = 0 if (headb01 == 9 | headb02 == 9 | headb03 == 9 | headb04 == 9 | headb05 == 9 | headb06 == 9 | headb07 == 9 | headb08 == 9 | headb09 == 9 | headb10 == 9 | headb11 == 9 | headb12 == 9 | headb13 == 9) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the third listed activity
replace headlsh = 1 if inlist(headb01,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb02,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb03,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb04,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb05,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb06,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb07,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb08,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb09,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb10,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb11,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb12,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0
replace headlsh = 1 if inlist(headb13,1,2,3,4,5,6,7,8,10,11,12,13,96) & headlsh!=0

* [b35] IADL: difficulty making telephone calls
replace headlph = 2 if headlph == 0
* Assign the number 0 if the participant reported difficulties performing the fourth listed activity
replace headlph = 0 if headlph == 1
* Assign the number 1 if the participant reported no difficulties performing the fourth listed activity
replace headlph = 1 if headlph == 2
* Assign the number 0 if the participant reported difficulties performing the fourth listed activity
replace headlph = 0 if (headb01 == 10 | headb02 == 10 | headb03 == 10 | headb04 == 10 | headb05 == 10 | headb06 == 10 | headb07 == 10 | headb08 == 10 | headb09 == 10 | headb10 == 10 | headb11 == 10 | headb12 == 10 | headb13 == 10) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the fourth listed activity
```

```
replace headlph = 1 if inlist(headb01,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb02,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb03,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb04,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb05,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb06,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb07,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb08,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb09,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb10,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb11,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb12,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0
replace headlph = 1 if inlist(headb13,1,2,3,4,5,6,7,8,9,11,12,13,96) & headlph!=0

* (1 omitted) IADL: difficulty talking medications
replace headlme = 2 if headlme == 0
* Assign the number 0 if the participant reported difficulties performing the fifth listed activity
replace headlme = 0 if headlme == 1
* Assign the number 1 if the participant reported no difficulties performing the fifth listed activity
replace headlme = 1 if headlme == 2
* Assign the number 0 if the participant reported difficulties performing the fifth listed activity
replace headlme = 0 if (headb01 == 11 | headb02 == 11 | headb03 == 11 | headb04 == 11 | headb05 == 11
 | headb06 == 11 | headb07 == 11 | headb08 == 11 | headb09 == 11 | headb10 == 11 | headb11 == 11 |
headb12 == 11 | headb13 == 11) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the fifth listed activity
replace headlme = 1 if inlist(headb01,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb02,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb03,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb04,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb05,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb06,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb07,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb08,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb09,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb10,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb11,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb12,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0
replace headlme = 1 if inlist(headb13,1,2,3,4,5,6,7,8,9,10,12,13,96) & headlme!=0

* [b36] IADL: difficulty doing work around the house or garden
replace headlho = 2 if headlho == 0
* Assign the number 0 if the participant reported difficulties performing the sixth listed activity
replace headlho = 0 if headlho == 1
* Assign the number 1 if the participant reported no difficulties performing the sixth listed activity
replace headlho = 1 if headlho == 2
* Assign the number 0 if the participant reported difficulties performing the sixth listed activity
replace headlho = 0 if (headb01 == 12 | headb02 == 12 | headb03 == 12 | headb04 == 12 | headb05 == 12
 | headb06 == 12 | headb07 == 12 | headb08 == 12 | headb09 == 12 | headb10 == 12 | headb11 == 12 |
headb12 == 12 | headb13 == 12) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the sixth listed activity
replace headlho = 1 if inlist(headb01,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb02,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb03,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb04,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb05,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb06,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb07,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb08,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb09,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb10,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb11,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb12,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
replace headlho = 1 if inlist(headb13,1,2,3,4,5,6,7,8,9,10,11,13,96) & headlho!=0
```

```
1053
1054  * (1 omitted) IADL: difficulty managing money, such as paying bills and keeping track of expenses
1055  replace headlmo = 2 if headlmo == 0
1056  * Assign the number 0 if the participant reported difficulties performing the seventh listed activity
1057  replace headlmo = 0 if headlmo == 1
1058  * Assign the number 1 if the participant reported no difficulties performing the seventh listed
      activity
1059  replace headlmo = 1 if headlmo == 2
1060  * Assign the number 0 if the participant reported difficulties performing the seventh listed activity
1061  replace headlmo = 0 if (headb01 == 13 | headb02 == 13 | headb03 == 13 | headb04 == 13 | headb05 == 13
       | headb06 == 13 | headb07 == 13 | headb08 == 13 | headb09 == 13 | headb10 == 13 | headb11 == 13 |
      headb12 == 13 | headb13 == 13) & wave==2
1062  * Assign the number 1 if the participant reported no difficulties performing the seventh listed
      activity
1063  replace headlmo = 1 if inlist(headb01,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1064  replace headlmo = 1 if inlist(headb02,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1065  replace headlmo = 1 if inlist(headb03,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1066  replace headlmo = 1 if inlist(headb04,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1067  replace headlmo = 1 if inlist(headb05,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1068  replace headlmo = 1 if inlist(headb06,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1069  replace headlmo = 1 if inlist(headb07,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1070  replace headlmo = 1 if inlist(headb08,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1071  replace headlmo = 1 if inlist(headb09,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1072  replace headlmo = 1 if inlist(headb10,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1073  replace headlmo = 1 if inlist(headb11,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1074  replace headlmo = 1 if inlist(headb12,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1075  replace headlmo = 1 if inlist(headb13,1,2,3,4,5,6,7,8,9,10,11,12,96) & headlmo!=0
1076
1077  * Save dataset with a new name
1078  save datavariables.dta
1079
1080  * WAVE 2, 3, 4, 5, 6, 7, 8, 9
1081  * [b16-b25] Mobility (reversed)
1082  * Replace variables as missing for any missing cases (coded as negative numbers in the ELSA dataset)
1083  replace hemobwa = . if hemobwa<0
1084  replace hemobsi = . if hemobsi<0
1085  replace hemobch = . if hemobch<0
1086  replace hemobcs = . if hemobcs<0
1087  replace hemobcl = . if hemobcl<0
1088  replace hemobst = . if hemobst<0
1089  replace hemobre = . if hemobre<0
1090  replace hemobpu = . if hemobpu<0
1091  replace hemobli = . if hemobli<0
1092  replace hemobpi = . if hemobpi<0
1093  replace hemob96 = . if hemob96<0
1094
1095  * [b16] Mobility: difficulty walking 100 yards
1096  replace hemobwa = 2 if hemobwa == 0
1097  * Assign the number 0 if the participant reported difficulties performing the first listed activity
1098  replace hemobwa = 0 if hemobwa == 1
1099  * Assign the number 1 if the participant reported no difficulties performing the first listed activity
1100  replace hemobwa = 1 if hemobwa == 2
1101  * Assign the number 0 if the participant reported difficulties performing the first listed activity
1102  replace hemobwa = 0 if (heada01 == 1 | heada02 == 1 | heada03 == 1 | heada04 == 1 | heada05 == 1 |
      heada06 == 1 | heada07 == 1 | heada08 == 1 | heada09 == 1 | heada10 == 1) & wave==2
1103  * Assign the number 1 if the participant reported no difficulties performing the first listed activity
1104  replace hemobwa = 1 if inlist(heada01,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0
1105  replace hemobwa = 1 if inlist(heada02,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0
1106  replace hemobwa = 1 if inlist(heada03,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0
1107  replace hemobwa = 1 if inlist(heada04,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0
1108  replace hemobwa = 1 if inlist(heada05,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0
1109  replace hemobwa = 1 if inlist(heada06,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0
1110  replace hemobwa = 1 if inlist(heada07,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0
```

```
replace hemobwa = 1 if inlist(heada08,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0
replace hemobwa = 1 if inlist(heada09,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0
replace hemobwa = 1 if inlist(heada10,2,3,4,5,6,7,8,9,10,96) & hemobwa!=0

* [b17] Mobility: difficulty sitting for about two hours
replace hemobsi = 2 if hemobsi == 0
* Assign the number 0 if the participant reported difficulties performing the second listed activity
replace hemobsi = 0 if hemobsi == 1
* Assign the number 1 if the participant reported no difficulties performing the second listed activity
replace hemobsi = 1 if hemobsi == 2
* Assign the number 0 if the participant reported difficulties performing the second listed activity
replace hemobsi = 0 if (heada01 == 2 | heada02 == 2 | heada03 == 2 | heada04 == 2 | heada05 == 2 | heada06 == 2 | heada07 == 2 | heada08 == 2 | heada09 == 2 | heada10 == 2) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the second listed activity
replace hemobsi = 1 if inlist(heada01,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0
replace hemobsi = 1 if inlist(heada02,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0
replace hemobsi = 1 if inlist(heada03,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0
replace hemobsi = 1 if inlist(heada04,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0
replace hemobsi = 1 if inlist(heada05,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0
replace hemobsi = 1 if inlist(heada06,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0
replace hemobsi = 1 if inlist(heada07,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0
replace hemobsi = 1 if inlist(heada08,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0
replace hemobsi = 1 if inlist(heada09,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0
replace hemobsi = 1 if inlist(heada10,1,3,4,5,6,7,8,9,10,96) & hemobsi!=0

* [b18] Mobility: difficulty getting up from a chair after sitting for long periods
replace hemobch = 2 if hemobch == 0
* Assign the number 0 if the participant reported difficulties performing the third listed activity
replace hemobch = 0 if hemobch == 1
* Assign the number 1 if the participant reported no difficulties performing the third listed activity
replace hemobch = 1 if hemobch == 2
* Assign the number 0 if the participant reported difficulties performing the third listed activity
replace hemobch = 0 if (heada01 == 3 | heada02 == 3 | heada03 == 3 | heada04 == 3 | heada05 == 3 | heada06 == 3 | heada07 == 3 | heada08 == 3 | heada09 == 3 | heada10 == 3) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the third listed activity
replace hemobch = 1 if inlist(heada01,1,2,4,5,6,7,8,9,10,96) & hemobch!=0
replace hemobch = 1 if inlist(heada02,1,2,4,5,6,7,8,9,10,96) & hemobch!=0
replace hemobch = 1 if inlist(heada03,1,2,4,5,6,7,8,9,10,96) & hemobch!=0
replace hemobch = 1 if inlist(heada04,1,2,4,5,6,7,8,9,10,96) & hemobch!=0
replace hemobch = 1 if inlist(heada05,1,2,4,5,6,7,8,9,10,96) & hemobch!=0
replace hemobch = 1 if inlist(heada06,1,2,4,5,6,7,8,9,10,96) & hemobch!=0
replace hemobch = 1 if inlist(heada07,1,2,4,5,6,7,8,9,10,96) & hemobch!=0
replace hemobch = 1 if inlist(heada08,1,2,4,5,6,7,8,9,10,96) & hemobch!=0
replace hemobch = 1 if inlist(heada09,1,2,4,5,6,7,8,9,10,96) & hemobch!=0
replace hemobch = 1 if inlist(heada10,1,2,4,5,6,7,8,9,10,96) & hemobch!=0

* [b19] Mobility: difficulty climbing several flights of stairs without resting
replace hemobcs = 2 if hemobcs == 0
* Assign the number 0 if the participant reported difficulties performing the fourth listed activity
replace hemobcs = 0 if hemobcs == 1
* Assign the number 1 if the participant reported no difficulties performing the fourth listed activity
replace hemobcs = 1 if hemobcs == 2
* Assign the number 0 if the participant reported difficulties performing the fourth listed activity
replace hemobcs = 0 if (heada01 == 4 | heada02 == 4 | heada03 == 4 | heada04 == 4 | heada05 == 4 | heada06 == 4 | heada07 == 4 | heada08 == 4 | heada09 == 4 | heada10 == 4) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the fourth listed activity
replace hemobcs = 1 if inlist(heada01,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0
replace hemobcs = 1 if inlist(heada02,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0
replace hemobcs = 1 if inlist(heada03,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0
```

```
replace hemobcs = 1 if inlist(heada04,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0
replace hemobcs = 1 if inlist(heada05,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0
replace hemobcs = 1 if inlist(heada06,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0
replace hemobcs = 1 if inlist(heada07,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0
replace hemobcs = 1 if inlist(heada08,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0
replace hemobcs = 1 if inlist(heada09,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0
replace hemobcs = 1 if inlist(heada10,1,2,3,5,6,7,8,9,10,96) & hemobcs!=0

* [b20] Mobility: difficulty climbing one flight of stairs without resting
replace hemobcl = 2 if hemobcl == 0
* Assign the number 0 if the participant reported difficulties performing the fifth listed activity
replace hemobcl = 0 if hemobcl == 1
* Assign the number 1 if the participant reported no difficulties performing the fifth listed activity
replace hemobcl = 1 if hemobcl == 2
* Assign the number 0 if the participant reported difficulties performing the fifth listed activity
replace hemobcl = 0 if (heada01 == 5 | heada02 == 5 | heada03 == 5 | heada04 == 5 | heada05 == 5 |
heada06 == 5 | heada07 == 5 | heada08 == 5 | heada09 == 5 | heada10 == 5) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the fifth listed activity
replace hemobcl = 1 if inlist(heada01,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0
replace hemobcl = 1 if inlist(heada02,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0
replace hemobcl = 1 if inlist(heada03,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0
replace hemobcl = 1 if inlist(heada04,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0
replace hemobcl = 1 if inlist(heada05,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0
replace hemobcl = 1 if inlist(heada06,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0
replace hemobcl = 1 if inlist(heada07,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0
replace hemobcl = 1 if inlist(heada08,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0
replace hemobcl = 1 if inlist(heada09,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0
replace hemobcl = 1 if inlist(heada10,1,2,3,4,6,7,8,9,10,96) & hemobcl!=0

* [b21] Mobility: difficulty stooping, kneeling, or crouching
replace hemobst = 2 if hemobst == 0
* Assign the number 0 if the participant reported difficulties performing the sixth listed activity
replace hemobst = 0 if hemobst == 1
* Assign the number 1 if the participant reported no difficulties performing the sixth listed activity
replace hemobst = 1 if hemobst == 2
* Assign the number 0 if the participant reported difficulties performing the sixth listed activity
replace hemobst = 0 if (heada01 == 6 | heada02 == 6 | heada03 == 6 | heada04 == 6 | heada05 == 6 |
heada06 == 6 | heada07 == 6 | heada08 == 6 | heada09 == 6 | heada10 == 6) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the sixth listed activity
replace hemobst = 1 if inlist(heada01,1,2,3,4,5,7,8,9,10,96) & hemobst!=0
replace hemobst = 1 if inlist(heada02,1,2,3,4,5,7,8,9,10,96) & hemobst!=0
replace hemobst = 1 if inlist(heada03,1,2,3,4,5,7,8,9,10,96) & hemobst!=0
replace hemobst = 1 if inlist(heada04,1,2,3,4,5,7,8,9,10,96) & hemobst!=0
replace hemobst = 1 if inlist(heada05,1,2,3,4,5,7,8,9,10,96) & hemobst!=0
replace hemobst = 1 if inlist(heada06,1,2,3,4,5,7,8,9,10,96) & hemobst!=0
replace hemobst = 1 if inlist(heada07,1,2,3,4,5,7,8,9,10,96) & hemobst!=0
replace hemobst = 1 if inlist(heada08,1,2,3,4,5,7,8,9,10,96) & hemobst!=0
replace hemobst = 1 if inlist(heada09,1,2,3,4,5,7,8,9,10,96) & hemobst!=0
replace hemobst = 1 if inlist(heada10,1,2,3,4,5,7,8,9,10,96) & hemobst!=0

* [b22] Mobility: difficulty reaching or extending arms above shoulder level
replace hemobre = 2 if hemobre == 0
* Assign the number 0 if the participant reported difficulties performing the seventh listed activity
replace hemobre = 0 if hemobre == 1
* Assign the number 1 if the participant reported no difficulties performing the seventh listed
activity
replace hemobre = 1 if hemobre == 2
* Assign the number 0 if the participant reported difficulties performing the seventh listed activity
replace hemobre = 0 if (heada01 == 7 | heada02 == 7 | heada03 == 7 | heada04 == 7 | heada05 == 7 |
heada06 == 7 | heada07 == 7 | heada08 == 7 | heada09 == 7 | heada10 == 7) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the seventh listed
activity
replace hemobre = 1 if inlist(heada01,1,2,3,4,5,6,8,9,10,96) & hemobre!=0
```

```
replace hemobre = 1 if inlist(heada02,1,2,3,4,5,6,8,9,10,96) & hemobre!=0
replace hemobre = 1 if inlist(heada03,1,2,3,4,5,6,8,9,10,96) & hemobre!=0
replace hemobre = 1 if inlist(heada04,1,2,3,4,5,6,8,9,10,96) & hemobre!=0
replace hemobre = 1 if inlist(heada05,1,2,3,4,5,6,8,9,10,96) & hemobre!=0
replace hemobre = 1 if inlist(heada06,1,2,3,4,5,6,8,9,10,96) & hemobre!=0
replace hemobre = 1 if inlist(heada07,1,2,3,4,5,6,8,9,10,96) & hemobre!=0
replace hemobre = 1 if inlist(heada08,1,2,3,4,5,6,8,9,10,96) & hemobre!=0
replace hemobre = 1 if inlist(heada09,1,2,3,4,5,6,8,9,10,96) & hemobre!=0
replace hemobre = 1 if inlist(heada10,1,2,3,4,5,6,8,9,10,96) & hemobre!=0

* [b23] Mobility: difficulty pulling or pushing large objects, like a living room chair
replace hemobpu = 2 if hemobpu == 0
* Assign the number 0 if the participant reported difficulties performing the eighth listed activity
replace hemobpu = 0 if hemobpu == 1
* Assign the number 1 if the participant reported no difficulties performing the eighth listed
activity
replace hemobpu = 1 if hemobpu == 2
* Assign the number 0 if the participant reported difficulties performing the eighth listed activity
replace hemobpu = 0 if (heada01 == 8 | heada02 == 8 | heada03 == 8 | heada04 == 8 | heada05 == 8 |
heada06 == 8 | heada07 == 8 | heada08 == 8 | heada09 == 8 | heada10 == 8) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the eighth listed
activity
replace hemobpu = 1 if inlist(heada01,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0
replace hemobpu = 1 if inlist(heada02,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0
replace hemobpu = 1 if inlist(heada03,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0
replace hemobpu = 1 if inlist(heada04,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0
replace hemobpu = 1 if inlist(heada05,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0
replace hemobpu = 1 if inlist(heada06,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0
replace hemobpu = 1 if inlist(heada07,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0
replace hemobpu = 1 if inlist(heada08,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0
replace hemobpu = 1 if inlist(heada09,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0
replace hemobpu = 1 if inlist(heada10,1,2,3,4,5,6,7,9,10,96) & hemobpu!=0

* [b24] Mobility: difficulty lifting or carrying weights over 10 pounds, like a heavy bag of groceries
replace hemobli = 2 if hemobli == 0
* Assign the number 0 if the participant reported difficulties performing the ninth listed activity
replace hemobli = 0 if hemobli == 1
* Assign the number 1 if the participant reported no difficulties performing the ninth listed activity
replace hemobli = 1 if hemobli == 2
* Assign the number 0 if the participant reported difficulties performing the ninth listed activity
replace hemobli = 0 if (heada01 == 9 | heada02 == 9 | heada03 == 9 | heada04 == 9 | heada05 == 9 |
heada06 == 9 | heada07 == 9 | heada08 == 9 | heada09 == 9 | heada10 == 9) & wave==2
* Assign the number 1 if the participant reported no difficulties performing the ninth listed activity
replace hemobli = 1 if inlist(heada01,1,2,3,4,5,6,7,8,10,96) & hemobli!=0
replace hemobli = 1 if inlist(heada02,1,2,3,4,5,6,7,8,10,96) & hemobli!=0
replace hemobli = 1 if inlist(heada03,1,2,3,4,5,6,7,8,10,96) & hemobli!=0
replace hemobli = 1 if inlist(heada04,1,2,3,4,5,6,7,8,10,96) & hemobli!=0
replace hemobli = 1 if inlist(heada05,1,2,3,4,5,6,7,8,10,96) & hemobli!=0
replace hemobli = 1 if inlist(heada06,1,2,3,4,5,6,7,8,10,96) & hemobli!=0
replace hemobli = 1 if inlist(heada07,1,2,3,4,5,6,7,8,10,96) & hemobli!=0
replace hemobli = 1 if inlist(heada08,1,2,3,4,5,6,7,8,10,96) & hemobli!=0
replace hemobli = 1 if inlist(heada09,1,2,3,4,5,6,7,8,10,96) & hemobli!=0
replace hemobli = 1 if inlist(heada10,1,2,3,4,5,6,7,8,10,96) & hemobli!=0

* [b25] Mobility: difficulty picking up a 5p coin from a table
replace hemobpi = 2 if hemobpi == 0
* Assign the number 0 if the participant reported difficulties performing the tenth listed activity
replace hemobpi = 0 if hemobpi == 1
* Assign the number 1 if the participant reported no difficulties performing the tenth listed activity
replace hemobpi = 1 if hemobpi == 2
* Assign the number 0 if the participant reported difficulties performing the tenth listed activity
replace hemobpi = 0 if (heada01 == 10 | heada02 == 10 | heada03 == 10 | heada04 == 10 | heada05 == 10
 | heada06 == 10 | heada07 == 10 | heada08 == 10 | heada09 == 10 | heada10 == 10) & wave==2
```

```
* Assign the number 1 if the participant reported no difficulties performing the tenth listed activity
replace hemobpi = 1 if inlist(heada01,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0
replace hemobpi = 1 if inlist(heada02,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0
replace hemobpi = 1 if inlist(heada03,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0
replace hemobpi = 1 if inlist(heada04,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0
replace hemobpi = 1 if inlist(heada05,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0
replace hemobpi = 1 if inlist(heada06,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0
replace hemobpi = 1 if inlist(heada07,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0
replace hemobpi = 1 if inlist(heada08,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0
replace hemobpi = 1 if inlist(heada09,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0
replace hemobpi = 1 if inlist(heada10,1,2,3,4,5,6,7,8,9,96) & hemobpi!=0

* Overwrite dataset, by replacing the previously saved file
save datavariables.dta, replace

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* Socio-economic covariate - Quintiles of BU total (non-pension) wealth
* Replace variable as missing for any missing cases (coded as negative numbers in the ELSA dataset)
replace totwq5_bu_s = . if totwq5_bu_s<0

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* [b41] Whether has self-reported limiting long-standing illness (reversed)
* Generate a new variable and assign the number 0 for participants with a limiting long-standing
illness
gen limiting = 0 if helim==1
* Assign the number 1 for participants with no long-standing illness or a long-standing illness that
is not limiting
replace limiting = 1 if heill == 2 | helim == 2

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* [b9-b15] (+ 1 omitted) Organisational memberships
* Replace variables as missing for any missing cases (coded as negative numbers in the ELSA dataset)
* [b9] Political party, trade union or environmental group
replace scorg01 = . if scorg01<0
* [b10] Tenants or resident group or neighbourhood watch
replace scorg02 = . if scorg02<0
* [b11] Member of a church or other religious group
replace scorg03 = . if scorg03<0
* [b12] Member of a charitable association
replace scorg04 = . if scorg04<0
* [b13] An education, arts or music group or evening class
replace scorg05 = . if scorg05<0
* (1 omitted) Member of a social club
replace scorg06 = . if scorg06<0
* [b14] Member of a sports club, gym, or exercise class
replace scorg07 = . if scorg07<0
* [b15] Member of any other organisations, clubs, or societies
replace scorg08 = . if scorg08<0

* Save dataset with a new name
save datavariables01.dta

* [b42-b43] (+ 1 omitted) Short Physical Performance Battery
* WAVE 2, 4, 6
* (1 omitted) Standing balance
tab mmbcsc
tab mmsssc
tab mmssre
tab mmsssc if mmssre==3
tab mmbcsc if mmsssc<0
tab mmsssc if mmssre<0
sum mmssti if mmssti!=-1
tab mmssre if mmssti!=-1
```

```
sum mmssna if mmssna!=-1
tab mmssna if mmssna!=-1 & wave==2
tab mmssna if mmssna!=-1 & inlist(wave,4,6)
tab mmsssc if mmssna!=-1

tab mmstsc
tab mmstsc if mmssna!=-1
tab mmssre if mmstsc==-1
tab mmstre
tab mmstsc if mmstre==3
sum mmstti if mmstti!=-1
tab mmstre if mmstti!=-1
sum mmstna if mmstna!=-1
tab mmstna if mmstna!=-1 & wave==2
tab mmstna if mmstna!=-1 & inlist(wave,4,6)
tab mmstsc if mmstna!=-1

tab mmftsc
tab mmftsc if mmstna!=-1
tab mmftsc if mmstti!=-1
tab mmftsc if mmssna!=-1 | mmssti!=-1
tab mmftsc if mmsssc<0
tab mmftre2
tab mmftsc if mmftre2==5
sum mmftti if mmftti!=-1
sum mmftti if inlist(mmftre2,2,4)
tab mmftre2 if mmftti!=-1
tab mmftre2 if inlist(mmftre2,2,4)
sum mmftti if mmftti >=3 & mmftti < 10
sum mmftti if mmftti < 3 & mmftti!=-1
tab mmftre2 if mmftti >=10 & mmftti!=.
sum mmftti if mmftti >=10 & mmftti!=.
sum mmftna if mmftna!=-1
tab mmftna if mmftna!=-1 & wave==2
tab mmftna if mmftna!=-1 & inlist(wave,4,6)
tab mmftsc if mmftna!=-1
tab mmftre2 if wave==6
tab mmftre

* Side-by-side stand
* Generate a new variable and assign the number 0 if the participant held for less than 10 seconds
or did not attempt the stand
gen sidebyside = 0 if inlist(mmssre,2,3)
* Assign the number 1 if the participant held for 10 seconds
replace sidebyside = 1 if mmssre==1

* Semi-tandem stand
* Generate a new variable and assign the number 0 if the participant scored 0 points in the
side-by-side stand
gen semitandem = 0 if sidebyside==0
* Assign the number 0 if the participant held for less than 10 seconds or did not attempt the stand
replace semitandem = 0 if inlist(mmstre,2,3)
* Assign the number 1 if the participant held for 10 seconds
replace semitandem = 1 if mmstre==1

* Full tandem stand
* Generate a new variable and assign the number 0 if the participant did not attempt the stand or
scored 0 points in the semi-tandem stand
gen tandem = 0 if (mmftre2 == 5 |semitandem == 0)
* Assign the number 2 if the participant held for at least 10 seconds, irrespective of age
replace tandem = 2 if (mmftre2 == 1 |mmftre2 == 3)
* Assign the number 1 if the participant held for at least 3 seconds but less than 10 seconds
replace tandem = 1 if mmftti >=3 & mmftti < 10
```

```
* Assign the number 0 if the participant held for less than 3 seconds and data are not missing
replace tandem = 0 if mmftti < 3 & mmftti!=-1
* Assign the number 2 if the participant held for at least 10 seconds (but less than 30 seconds) and was aged less than 70 years
replace tandem = 2 if mmftti >=10 & mmftti!=. & mmftre2==4

* Generate a new variable equal to the sum of the points scored on the side-by-side, semi-tandem, and full tandem stands (three items) to create a total balance score (range 0-4)
gen balance = sidebyside + semitandem + tandem

gen balance2 = 0 if inlist(mmssre,2,3)
replace balance2 = 1 if mmssre==1 & inlist(mmstre,2,3)
replace balance2 = 2 if mmstre==1 & tandem==0
replace balance2 = 3 if mmstre==1 & tandem==1
replace balance2 = 4 if mmstre==1 & tandem==2

* WAVE 2, 4, 6
* [b42] Repeated chair stands
tab mmcrav
* Generate a new variable and assign a missing value if there was no suitable chair available or data are missing
gen repcstest = . if mmcrav==2 | mmcrav<0
tab mmcrsc
tab mmcrsc if mmcrav==2 | mmcrav<0
tab mmcrre
tab mmcrre if mmcrav==2 | mmcrav<0 | inlist(mmcrsc,-8,2)
* Assign the number 0 if the participant did not feel it was safe to attempt the single chair rise or the response was coded "Don't know"
replace repcstest = 0 if inlist(mmcrsc,-8,2)
* Assign the number 0 if the participant used their arms to stand in the single chair rise or did not complete the test
replace repcstest = 0 if inlist(mmcrre,2,3)
tab mmcrav if mmcrre==-1
tab mmcrsc if mmcrre==-1
sum mmcrna if mmcrna!=-1
tab mmcrna if mmcrna!=-1 & wave==2
tab mmcrna if mmcrna!=-1 & inlist(wave,4,6)
tab mmrrsc
tab mmcrre if mmrrsc==-1
tab mmrrre
tab mmrrsc if mmrrre==-1
tab mmrrre if inlist(mmrrsc,2,-1)
* Assign the number 0 if the participant did not feel it was safe to attempt multiple chair rises (and subsequently did not perform the multiple chair rise test)
replace repcstest = 0 if mmrrsc==2 & mmrrre==-1
* Assign the number 0 if the participant completed less than five sit-to-stands
replace repcstest = 0 if inlist(mmrrre,0,1,2,3,4)
tab mmrrfti if mmrrfti<0
tab mmrrre if mmrrfti<0
sum mmrrfti if mmrrre>=5
sum mmrrfti if mmrrfti>=0
tab mmrrfti if mmrrfti<0 & wave==2
tab mmrrfti if mmrrfti<0 & wave==4
tab mmrrfti if mmrrfti<0 & wave==6
* Assign a missing value if the participant completed five or more sit-to-stands but their time to complete five rises was coded as "Don't know" or the test was not timed correctly
replace repcstest = . if inlist(mmrrfti,-8,-3) & mmrrre>=5
tab mmrrfti if mmrrfti>=0 & mmrrfti<4
* Assign a missing value if the participant completed five or more sit-to-stands but their time to complete five rises was equal to 0 or 1 seconds (i.e., outlier)
replace repcstest = . if inlist(mmrrfti,0,1) & inlist(mmrrre,5,6,7,8,9,10)
sum mmrrfti if mmrrfti<=11.19 & mmrrfti>=0
sum mmrrfti if mmrrfti<=11.19 & mmrrfti>1
```

```
* Assign the number 4 if the participant completed five sit-to-stands in less than or equal to 11.19 seconds and it took them more than 1 second
replace repcstest = 4 if mmrrfti <= 11.19 & mmrrfti > 1 & inlist(mmrrre,5,6,7,8,9,10)
sum mmrrfti if mmrrfti>=16.7 & mmrrfti<=60
* Assign the number 1 if the participant completed five sit-to-stands in 16.7 seconds or more but less than 60 seconds
replace repcstest = 1 if mmrrfti >= 16.7 & mmrrfti <= 60 & inlist(mmrrre,5,6,7,8,9,10)
sum mmrrfti if mmrrfti>=13.7 & mmrrfti<16.7
* Assign the number 2 if the participant completed five sit-to-stands in 13.7 seconds or more but less than 16.7 seconds
replace repcstest = 2 if mmrrfti >= 13.7 & mmrrfti < 16.7 & inlist(mmrrre,5,6,7,8,9,10)
sum mmrrfti if mmrrfti>=11.2 & mmrrfti<13.7
* Assign the number 3 if the participant completed five sit-to-stands in 11.2 seconds or more but less than 13.7 seconds
replace repcstest = 3 if mmrrfti >= 11.2 & mmrrfti < 13.7 & inlist(mmrrre,5,6,7,8,9,10)
sum mmrrfti if mmrrfti > 60
* Assign the number 0 if the participant completed five sit-to-stands in more than 60 seconds and data are not missing
replace repcstest = 0 if mmrrfti > 60 & mmrrfti <= 64 & inlist(mmrrre,5,6,7,8,9,10)

tab mmrrre if inlist(mmrroc,1,3)
tab mmrrre if mmrroc==2
tab mmrrre if mmrroc==4
tab mmrrre if inlist(mmrroc,1,2,3,4)

sum mmrrna if mmrrna!=-1
tab mmrrna if mmrrna!=-1 & wave==2
tab mmrrna if mmrrna!=-1 & inlist(wave,4,6)
tab mmrrre if mmrrna!=-1 & inlist(mmrroc,3,4)
tab mmrrre if mmrrna!=-1 & inlist(mmrroc,1,2)
tab mmrrsc if mmrrna!=-1
tab mmrrna if mmrrre==5 & inlist(mmrroc,1,2)
sum mmrrfti if mmrrre==5 & inlist(mmrroc,1,2) & mmrrna!=-1

* Save dataset with a new name
save datavariables02.dta

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* [b43] Gait speed
tab mmschs if indager >=60
tab mmalone if indager >=60
tab mmschs if mmalone==-1 & indager >=60
tab mmschs if mmalone==-9 & indager >=60
tab mmschs if mmalone==-8 & indager >=60
tab mmschs if mmalone==-2 & indager >=60
tab mmschs if mmalone==1 & indager >=60
tab mmschs if mmalone==2 & indager >=60
tab mmschs if mmalone==3 & indager >=60
tab mmhss if indager >=60
tab mmalone if mmhss==-1 & indager >=60
tab mmalone if mmhss==-9 & indager >=60
tab mmalone if mmhss==-8 & indager >=60
tab mmalone if mmhss==-2 & indager >=60
tab mmalone if mmhss==1 & indager >=60
tab mmalone if mmhss==2 & indager >=60
tab mmalone if mmhss==3 & indager >=60
tab mmalone if mmhss==4 & indager >=60
tab mmwill if indager >=60
tab mmhss if mmwill==-9 & indager >=60
tab mmhss if mmwill==-8 & indager >=60
tab mmhss if mmwill==-2 & indager >=60
tab mmhss if mmwill==-1 & indager >=60
tab mmhss if mmwill==1 & indager >=60
```

```
tab mmhss if mmwill==2 & indager >=60
tab mmsaf if indager >=60
tab mmwill if mmsaf==1 & indager >=60
tab mmwill if mmsaf==2 & indager >=60
tab mmavsp if indager >=60
tab mmsaf if mmavsp==-9 & indager >=60
tab mmsaf if mmavsp==-8 & indager >=60
tab mmsaf if mmavsp==-2 & indager >=60
tab mmsaf if mmavsp==-1 & indager >=60
tab mmsaf if mmavsp==1 & indager >=60
tab mmsaf if mmavsp==2 & indager >=60
tab mmwala if indager >=60
tab mmavsp if inlist(mmwala,1,2) & indager >=60

tab mmtrya if indager >=60
tab mmwala if inlist(mmtrya,-9,1,2,3) & indager >=60
tab mmwala if mmtrya==4 & indager >=60
tab mmwala if mmtrya==-8 & indager >=60
tab mmwala if mmtrya==-2 & indager >=60
tab mmwlka if mmwlka<0 & indager >=60
sum mmwlka if mmwlka>=0 & indager >=60
tab mmtrya if mmwlka==-1 & indager >=60
tab mmtrya if mmwlka==-2 & indager >=60
tab mmtrya if mmwlka==-8 & indager >=60
tab mmtrya if mmwlka==-9 & indager >=60
tab mmtrya if mmwlka>=0 & indager >=60

tab mmtryb if indager >=60
tab mmtrya if inlist(mmtryb,-8,1,2,3,4) & indager >=60
tab mmwlka if inlist(mmtryb,-9,-8,-2,-1) & indager >=60
tab mmwlkb if mmwlkb<0 & indager >=60
sum mmwlkb if mmwlkb>=0 & indager >=60
tab mmtryb if mmwlkb==-1 & indager >=60
tab mmtryb if mmwlkb==-2 & indager >=60
tab mmtryb if mmwlkb==-8 & indager >=60
tab mmtryb if mmwlkb==-9 & indager >=60
tab mmtryb if mmwlkb>=0 & indager >=60

tab mmwlka if mmwlka>=0 & mmwlka<2 & indager >=60
tab mmwlkb if mmwlkb>=0 & mmwlkb<2 & indager >=60

sum idauniq if ((mmwlka>=0 & mmwlka!=.) | (mmwlkb>=0 & mmwlkb!=.)) & indager >=60
sum idauniq if mmwlka>=0 & mmwlkb>=0 & mmwlka!=. & mmwlkb!=. & indager >=60

sum idauniq if mmwlka<0 & indager >=60
keep if mmwlka<0 & indager >=60
tab mmschs if indager >=60
tab mmalone if indager >=60
tab mmschs if mmalone==-1 & indager >=60
tab mmschs if mmalone==3 & indager >=60
tab mmschs if mmalone==-9 & indager >=60
tab mmschs if mmalone==-8 & indager >=60
tab mmschs if mmalone==-2 & indager >=60
tab mmhss if indager >=60
tab mmalone if mmhss==-1 & indager >=60
tab mmwill if indager >=60
tab mmhss if mmwill==-1 & indager >=60
tab mmsaf if indager >=60
tab mmwill if mmsaf==-1 & indager >=60
tab mmavsp if indager >=60
tab mmsaf if mmavsp==-1 & indager >=60
tab mmwala if indager >=60
tab mmavsp if mmwala==-1 & indager >=60
```

```
tab mmtrya if indager >=60
tab mmwala if mmtrya==-1 & indager >=60
tab mmwlka if indager >=60
tab mmtrya if mmwlka==-1 & indager >=60
tab mmtrya if mmwlka<0 & mmwlka!=-1 & indager >=60

clear
* Use full dataset
use datavariables02.dta

* Generate a new variable duplicating the "time taken for first walk" variable if data are not missing
gen walk1 = mmwlka if mmwlka>=0 & indager >=60
* Generate a new variable duplicating the "time taken for second walk" variable if data are not missing
gen walk2 = mmwlkb if mmwlkb>=0 & indager >=60
* Generate a new variable equal to the fastest time of the two walks (or the only time available if only one attempt was performed or recorded)
egen gaittime = rowmin(walk1 walk2) if ((mmwlka>=0 & mmwlka!=.) | (mmwlkb>=0 & mmwlkb!=.)) & indager >=60
sum gaittime
sum gaittime if (mmwlka>=0 & mmwlka!=.) & (mmwlkb<0|mmwlkb==.)
sum mmwlka if (mmwlka>=0 & mmwlka!=.) & (mmwlkb<0|mmwlkb==.)

* Generate a new variable
gen gaittest = .
* Assign the number 1 if the participant completed the gait test in more than or equal to 5.7 seconds
replace gaittest = 1 if gaittime >= 5.7 & gaittime!=.
* Assign the number 2 if the participant completed the gait test in more than or equal to 4.1 seconds and less than 5.7 seconds
replace gaittest = 2 if gaittime >= 4.1 & gaittime < 5.7
* Assign the number 3 if the participant completed the gait test in more than or equal to 3.2 seconds and less than 4.1 seconds
replace gaittest = 3 if gaittime >= 3.2 & gaittime < 4.1
* Assign the number 4 if the participant completed the gait test in less than 3.2 seconds
replace guralnik = 4 if gaittime < 3.2
* Assign the number 0 if a) the participant was not able to walk alone (with aid); b) a health condition (i.e., recent surgery, injury, other health condition) prevented the participant from walking; c) the interviewer felt it was not safe to continue the test; d) the respondent did not feel the walk would be safe; or e) the participant attempted the walk but was unable to complete it or was stopped by the interviewer because of safety reasons
replace guralnik = 0 if (mmalone==3 | inlist(mmhss,2,3,4) | mmsaf==2  | mmwala==2 | inlist(mmtrya,2,3)) & indager >=60

* Save dataset with a new name
save datavariables03.dta

* WAVE 2, 4, 5, 6, 7, 8, 9
* [b44] Self-reported general health (reversed)
* Replace variable as missing for any missing cases (coded as negative numbers in the ELSA dataset)
replace hehelf = . if hehelf<0
* Reverse the self-rated health variable
revrs hehelf

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* [b1-b6] (+ 2 omitted) Depressive symptoms
* Replace variables as missing for any missing cases (coded as negative numbers in the ELSA dataset)
* [b1] Whether respondent has felt depressed much of the time during the past week
replace psceda = . if psceda<0
* (1 omitted) Whether respondent felt everything they did during the past week was an effort
replace pscedb = . if pscedb<0
* [b2] Whether respondent felt their sleep was restless during the past week
replace pscedc = . if pscedc<0
* [b3] Whether respondent was happy much of the time during the past week (reversed)
```

```
replace pscedd = . if pscedd<0
* [b4] Whether respondent felt lonely must of the time during the past week
replace pscede = . if pscede<0
* [b5] Whether respondent enjoyed life much of the time during the past week (reversed)
replace pscedf = . if pscedf<0
* [b6] Whether respondent felt sad much of the time during the past week
replace pscedg = . if pscedg<0
* (1 omitted) Whether respondent could not get going much of the time during the past week
replace pscedh = . if pscedh<0

* Recode to the number 0 if participant answered "Yes"
replace psceda = 0 if psceda == 1
* Recode to the number 1 if participant answered "No"
replace psceda = 1 if psceda == 2

* Recode to the number 0 if participant answered "Yes"
replace pscedb = 0 if pscedb == 1
* Recode to the number 1 if participant answered "No"
replace pscedb = 1 if pscedb == 2

* Recode to the number 0 if participant answered "Yes"
replace pscedc = 0 if pscedc == 1
* Recode to the number 1 if participant answered "No"
replace pscedc = 1 if pscedc == 2

* Recode to the number 0 if participant answered "Yes"
replace pscede = 0 if pscede == 1
* Recode to the number 1 if participant answered "No"
replace pscede = 1 if pscede == 2

* Recode to the number 0 if participant answered "Yes"
replace pscedg = 0 if pscedg == 1
* Recode to the number 1 if participant answered "No"
replace pscedg = 1 if pscedg == 2

* Recode to the number 0 if participant answered "Yes"
replace pscedh = 0 if pscedh == 1
* Recode to the number 1 if participant answered "No"
replace pscedh = 1 if pscedh == 2

* Recode to the number 0 if participant answered "No"
replace pscedd = 0 if pscedd == 2
* Recode to the number 0 if participant answered "No"
replace pscedf = 0 if pscedf == 2

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* (3 omitted) Loneliness (reversed)
* Replace variables as missing for any missing cases (coded as negative numbers in the ELSA dataset)
* (1 omitted) How often respondent feels they lack companionship
replace scfeela = . if scfeela<0
* (1 omitted) How often respondent feels left out
replace scfeelb = . if scfeelb<0
* (1 omitted) How often respondent feels isolated from others
replace scfeelc = . if scfeelc<0

* Reverse the negatively framed variables
revrs scfeela
revrs scfeelb
revrs scfeelc

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* Outcome variable - Quality-of-life
* Replace variables as missing for any missing cases (coded as negative numbers in the ELSA dataset)
```

```
replace scqola = . if scqola<0
replace scqolb = . if scqolb<0
replace scqolc = . if scqolc<0
replace scqold = . if scqold<0
replace scqole = . if scqole<0
replace scqolf = . if scqolf<0
replace scqolg = . if scqolg<0
replace scqolh = . if scqolh<0
replace scqoli = . if scqoli<0
replace scqolj = . if scqolj<0
replace scqolk = . if scqolk<0
replace scqoll = . if scqoll<0
replace scqolm = . if scqolm<0
replace scqoln = . if scqoln<0
replace scqolo = . if scqolo<0
replace scqolp = . if scqolp<0
replace scqolq = . if scqolq<0
replace scqolr = . if scqolr<0
replace scqols = . if scqols<0

* Recode each item into a 0-3 scale (13 of the 19 items were reversed)
replace scqola = 0 if scqola == 1
replace scqola = 1 if scqola == 2
replace scqola = 2 if scqola == 3
replace scqola = 3 if scqola == 4
replace scqolb = 0 if scqolb == 1
replace scqolb = 1 if scqolb == 2
replace scqolb = 2 if scqolb == 3
replace scqolb = 3 if scqolb == 4
revrs scqolc
replace revscqolc = 0 if revscqolc == 1
replace revscqolc = 1 if revscqolc == 2
replace revscqolc = 2 if revscqolc == 3
replace revscqolc = 3 if revscqolc == 4
replace scqold = 0 if scqold == 1
replace scqold = 1 if scqold == 2
replace scqold = 2 if scqold == 3
replace scqold = 3 if scqold == 4
revrs scqole
replace revscqole = 0 if revscqole == 1
replace revscqole = 1 if revscqole == 2
replace revscqole = 2 if revscqole == 3
replace revscqole = 3 if revscqole == 4
replace scqolf = 0 if scqolf == 1
replace scqolf = 1 if scqolf == 2
replace scqolf = 2 if scqolf == 3
replace scqolf = 3 if scqolf == 4
revrs scqolg
replace revscqolg = 0 if revscqolg == 1
replace revscqolg = 1 if revscqolg == 2
replace revscqolg = 2 if revscqolg == 3
replace revscqolg = 3 if revscqolg == 4
replace scqolh = 0 if scqolh == 1
replace scqolh = 1 if scqolh == 2
replace scqolh = 2 if scqolh == 3
replace scqolh = 3 if scqolh == 4
replace scqoli = 0 if scqoli == 1
replace scqoli = 1 if scqoli == 2
replace scqoli = 2 if scqoli == 3
replace scqoli = 3 if scqoli == 4
revrs scqolj
replace revscqolj = 0 if revscqolj == 1
replace revscqolj = 1 if revscqolj == 2
```

```
1759  replace revscqolj = 2 if revscqolj == 3
1760  replace revscqolj = 3 if revscqolj == 4
1761  revrs scqolk
1762  replace revscqolk = 0 if revscqolk == 1
1763  replace revscqolk = 1 if revscqolk == 2
1764  replace revscqolk = 2 if revscqolk == 3
1765  replace revscqolk = 3 if revscqolk == 4
1766  revrs scqoll
1767  replace revscqoll = 0 if revscqoll == 1
1768  replace revscqoll = 1 if revscqoll == 2
1769  replace revscqoll = 2 if revscqoll == 3
1770  replace revscqoll = 3 if revscqoll == 4
1771  revrs scqolm
1772  replace revscqolm = 0 if revscqolm == 1
1773  replace revscqolm = 1 if revscqolm == 2
1774  replace revscqolm = 2 if revscqolm == 3
1775  replace revscqolm = 3 if revscqolm == 4
1776  revrs scqoln
1777  replace revscqoln = 0 if revscqoln == 1
1778  replace revscqoln = 1 if revscqoln == 2
1779  replace revscqoln = 2 if revscqoln == 3
1780  replace revscqoln = 3 if revscqoln == 4
1781  revrs scqolo
1782  replace revscqolo = 0 if revscqolo == 1
1783  replace revscqolo = 1 if revscqolo == 2
1784  replace revscqolo = 2 if revscqolo == 3
1785  replace revscqolo = 3 if revscqolo == 4
1786  revrs scqolp
1787  replace revscqolp = 0 if revscqolp == 1
1788  replace revscqolp = 1 if revscqolp == 2
1789  replace revscqolp = 2 if revscqolp == 3
1790  replace revscqolp = 3 if revscqolp == 4
1791  revrs scqolq
1792  replace revscqolq = 0 if revscqolq == 1
1793  replace revscqolq = 1 if revscqolq == 2
1794  replace revscqolq = 2 if revscqolq == 3
1795  replace revscqolq = 3 if revscqolq == 4
1796  revrs scqolr
1797  replace revscqolr = 0 if revscqolr == 1
1798  replace revscqolr = 1 if revscqolr == 2
1799  replace revscqolr = 2 if revscqolr == 3
1800  replace revscqolr = 3 if revscqolr == 4
1801  revrs scqols
1802  replace revscqols = 0 if revscqols == 1
1803  replace revscqols = 1 if revscqols == 2
1804  replace revscqols = 2 if revscqols == 3
1805  replace revscqols = 3 if revscqols == 4
1806
1807  * Generate a new variable equal to the sum of the 19 items (range 0-57)
1808  gen QoL = scqola + scqolb + revscqolc + scqold + revscqole + scqolf + revscqolg + scqolh + scqoli +
      revscqolj + revscqolk + revscqoll + revscqolm + revscqoln + revscqolo + revscqolp + revscqolq +
      revscqolr + revscqols
1809
1810  * Overwrite dataset, by replacing the previously saved file
1811  save datavariables03.dta, replace
1812
1813  * WAVE 2, 3, 4, 5, 6, 7, 8, 9
1814  * Independent variable - Alcohol consumption
1815  * Replace variable as missing for any missing cases (coded as negative numbers in the ELSA dataset)
1816  replace scako = . if scako<0
1817  * Assign the number 0 if the participant reported having an alcoholic drink once a month or less
      during the last 12 month
1818  replace scako = 0 if inlist(scako,5,6,7,8)
```

```
* Assign the number 2 if the participant reported having an alcoholic drink a) almost every day; or b) five or six days a week
replace scako = 2 if scako==1
* Assign the number 1 if the participant reported having an alcoholic drink a) three or four days a week; or b) once or twice a week
replace scako = 1 if inlist(scako,3,4)

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* Socio-economic covariate - Education
* Generate a new variable
gen education = .
* Assign the number 0 if the participant does not have any formal qualifications
replace education = 0 if edqual==7
* Assign the number 1 if the participant has A level equivalent, O level equivalent, or other grade equivalent
replace education = 1 if inlist(edqual,3,4,5)
* Assign the number 2 if the participant has completed some higher education (below degree), or has a degree or equivalent
replace education = 2 if inlist(edqual,1,2)

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* Demographic covariate - Ethnicity
* Replace variable as missing for any missing cases (coded as negative numbers in the ELSA dataset)
replace fqethnr = . if fqethnr<0
* Assign the number 0 if the participant is White
replace fqethnr = 0 if fqethnr==1
* Assign the number 1 if the participant is Non-White
replace fqethnr = 1 if fqethnr==2

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* Demographic covariate - Biological sex
* Replace variable as missing for any missing cases (coded as negative numbers in the ELSA dataset)
replace disex = . if disex<0
* Assign the number 0 if the participant is male
replace disex = 0 if disex==1
* Assign the number 1 if the participant is female
replace disex = 1 if disex==2

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* Independent variable - Physical activity
* Generate a new variable
gen activity2 = .
* Assign the number 3 if the participant partakes in vigorous activity more than once a week or once a week
replace activity2 = 3 if heacta==1 | heacta==2
* Assign the number 2 if the participant partakes in moderate activity more than once a week or once a week, and takes part in vigorous activity less than once a week
replace activity2 = 2 if (heactb==1 | heactb==2) & inlist(heacta,3,4)
* Assign the number 1 if the participant partakes in mild activity more than once a week or once a week, and takes part in moderate and vigorous activities less than once a week
replace activity2 = 1 if (heactc==1 | heactc==2) & inlist(heacta,3,4) & inlist(heactb,3,4)
* Assign the number 0 if the participant does not take part in activity of any intensity once a week or more
replace activity2 = 0 if inlist(heacta,3,4) & inlist(heactb,3,4) & inlist(heactc,3,4)
* Replace the variable as missing for participants with missing cases on all three variables
replace activity2 = . if inlist(heacta,.) & inlist(heactb,.) & inlist(heactc,.)

* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* Independent variable - Smoking status
* Generate a new variable duplicating the heske (reason disputed reported smoking from previous wave) variable at Wave 2
gen heske2 = heske if wave==2
* Declare a panel dataset with participant ID "idauniq" and time variable "wave"
```

```
1873  tsset idauniq wave
1874  * Generate a completely balanced dataset (i.e., all participants have a row for each wave from 2 to 9)
1875  tsfill, full
1876  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 2 to the
      follow-up waves) by participant ID
1877  bysort idauniq: carryforward heske2, replace
1878
1879  * Generate a new variable duplicating the heske (reason disputed reported smoking from previous
      wave) variable at Wave 3
1880  gen heske3 = heske if wave==3
1881  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
1882  tsset idauniq wave
1883  * Generate a completely balanced dataset
1884  tsfill, full
1885  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 3 to the
      follow-up waves) by participant ID
1886  bysort idauniq: carryforward heske3, replace
1887  * Sort data in memory by ascending values of the participant ID variable "idauniq" and descending
      values of the time variable "wave"
1888  gsort idauniq - wave
1889  * Carryforward (in a backward way) observations with respect to the time variable "wave" by
      participant ID
1890  bysort idauniq: carryforward heske3, replace
1891
1892  * Generate a new variable duplicating the heske (reason disputed reported smoking from previous
      wave) variable at Wave 4
1893  gen heske4 = heske if wave==4
1894  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
1895  tsset idauniq wave
1896  * Generate a completely balanced dataset
1897  tsfill, full
1898  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 4 to the
      follow-up waves) by participant ID
1899  bysort idauniq: carryforward heske4, replace
1900  * Sort data in memory by ascending values of the participant ID variable "idauniq" and descending
      values of the time variable "wave"
1901  gsort idauniq - wave
1902  * Carryforward (in a backward way) observations with respect to the time variable "wave" by
      participant ID
1903  bysort idauniq: carryforward heske4, replace
1904
1905  * Generate a new variable duplicating the heske (reason disputed reported smoking from previous
      wave) variable at Wave 5
1906  gen heske5 = heske if wave==5
1907  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
1908  tsset idauniq wave
1909  * Generate a completely balanced dataset
1910  tsfill, full
1911  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 5 to the
      follow-up waves) by participant ID
1912  bysort idauniq: carryforward heske5, replace
1913  * Sort data in memory by ascending values of the participant ID variable "idauniq" and descending
      values of the time variable "wave"
1914  gsort idauniq - wave
1915  * Carryforward (in a backward way) observations with respect to the time variable "wave" by
      participant ID
1916  bysort idauniq: carryforward heske5, replace
1917
1918  * Generate a new variable duplicating the heske (reason disputed reported smoking from previous
      wave) variable at Wave 6
1919  gen heske6 = heske if wave==6
1920  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
1921  tsset idauniq wave
```

```
* Generate a completely balanced dataset
tsfill, full
* Carryforward observations with respect to the time variable "wave" (i.e., from Wave 6 to the follow-up waves) by participant ID
bysort idauniq: carryforward heske6, replace
* Sort data in memory by ascending values of the participant ID variable "idauniq" and descending values of the time variable "wave"
gsort idauniq - wave
* Carryforward (in a backward way) observations with respect to the time variable "wave" by participant ID
bysort idauniq: carryforward heske6, replace

* Generate a new variable duplicating the heske (reason disputed reported smoking from previous wave) variable at Wave 7
gen heske7 = heske if wave==7
* Declare a panel dataset with participant ID "idauniq" and time variable "wave"
tsset idauniq wave
* Generate a completely balanced dataset
tsfill, full
* Carryforward observations with respect to the time variable "wave" (i.e., from Wave 7 to the follow-up waves) by participant ID
bysort idauniq: carryforward heske7, replace
* Sort data in memory by ascending values of the participant ID variable "idauniq" and descending values of the time variable "wave"
gsort idauniq - wave
* Carryforward (in a backward way) observations with respect to the time variable "wave" by participant ID
bysort idauniq: carryforward heske7, replace

* Generate a new variable duplicating the heske (reason disputed reported smoking from previous wave) variable at Wave 8
gen heske8 = heske if wave==8
* Declare a panel dataset with participant ID "idauniq" and time variable "wave"
tsset idauniq wave
* Generate a completely balanced dataset
tsfill, full
* Carryforward observations with respect to the time variable "wave" (i.e., from Wave 8 to the follow-up wave) by participant ID
bysort idauniq: carryforward heske8, replace
* Sort data in memory by ascending values of the participant ID variable "idauniq" and descending values of the time variable "wave"
gsort idauniq - wave
* Carryforward (in a backward way) observations with respect to the time variable "wave" by participant ID
bysort idauniq: carryforward heske8, replace

* Generate a new variable duplicating the heske (reason disputed reported smoking from previous wave) variable at Wave 9
gen heske9 = heske if wave==9
* Declare a panel dataset with participant ID "idauniq" and time variable "wave"
tsset idauniq wave
* Generate a completely balanced dataset
tsfill, full
* Sort data in memory by ascending values of the participant ID variable "idauniq" and descending values of the time variable "wave"
gsort idauniq - wave
* Carryforward (in a backward way) observations with respect to the time variable "wave" by participant ID
bysort idauniq: carryforward heske9, replace

* Generate a new variable duplicating the bhesmkc (whether reported ever smoked cigarettes at Wave 1) variable at Wave 2
gen bhesmkc = bhesmk
```

```
1970  * Assign the number 1 if the participant reported no longer smoking cigarettes by last interview
      (Wave 1)
1971  replace bhesmkc = 1 if heske==2 & wave==2
1972  * Assign the number 2 if the participant reported they
1973  replace bhesmkc = 2 if heske==1 & wave==2
1974
1975  * Generate a new variable and assign the number 0 if the participant reported never having smoked
      cigarettes
1976  gen smoking = 0 if hesmk==2
1977  * Assign the number 0 if the participant reported never having smoked cigarettes at Wave 1 and
      reported that they do not smoke cigarettes at all nowadays
1978  replace smoking = 0 if bhesmkc==2 & heska==2
1979  * Assign the number 1 if the participant reported having ever smoked cigarettes but reported that
      they do not smoke cigarettes at all nowadays
1980  replace smoking = 1 if (hesmk==1 | bhesmkc==1) & heska==2
1981  * Assign the number 2 if the participant reported smoking nowadays
1982  replace smoking = 2 if heska==1
1983
1984  * Assign the number 1 if the participant reported that they stopped smoking between Wave 1 and Wave
      2 and reported that they do not smoke cigarettes at all nowadays at Wave 2
1985  replace smoking = 1 if heske==3 & heska==2 & wave==2
1986  tab heskd
1987  * Assign the number 0 if the participant reported never having smoked cigarettes and reported that
      they do not smoke cigarettes at all nowadays at Wave 2
1988  replace smoking = 0 if heske3==1 & heska==2 & wave==2
1989  * Assign the number 1 if the participant reported that they were no longer smoking cigarettes by
      Wave 2 and reported that they do not smoke cigarettes at all nowadays at Wave 2
1990  replace smoking = 1 if heske3==2 & heska==2 & wave==2
1991
1992  * Assign the number 1 if the participant reported that they stopped smoking between Wave 2 and Wave
      3 and reported that they do not smoke cigarettes at all nowadays at Wave 3
1993  replace smoking = 1 if heske==3 & heska==2 & wave==3
1994  * Assign the number 0 if the participant reported never having smoked cigarettes and reported that
      they do not smoke cigarettes at all nowadays at Wave 3
1995  replace smoking = 0 if heske4==1 & heska==2 & wave==3
1996  * Assign the number 1 if the participant reported that they were no longer smoking cigarettes by
      Wave 3 and reported that they do not smoke cigarettes at all nowadays at Wave 3
1997  replace smoking = 1 if heske4==2 & heska==2 & wave==3
1998
1999  * Assign the number 1 if the participant reported that they stopped smoking between Wave 3 and Wave
      4 and reported that they do not smoke cigarettes at all nowadays at Wave 4
2000  replace smoking = 1 if heske==3 & heska==2 & wave==4
2001  * Assign the number 0 if the participant reported never having smoked cigarettes and reported that
      they do not smoke cigarettes at all nowadays at Wave 4
2002  replace smoking = 0 if heske5==1 & heska==2 & wave==4
2003  * Assign the number 1 if the participant reported that they were no longer smoking cigarettes by
      Wave 4 and reported that they do not smoke cigarettes at all nowadays at Wave 4
2004  replace smoking = 1 if heske5==2 & heska==2 & wave==4
2005
2006  * Assign the number 1 if the participant reported that they stopped smoking between Wave 4 and Wave
      5 and reported that they do not smoke cigarettes at all nowadays at Wave 5
2007  replace smoking = 1 if heske==3 & heska==2 & wave==5
2008  * Assign the number 0 if the participant reported never having smoked cigarettes and reported that
      they do not smoke cigarettes at all nowadays at Wave 5
2009  replace smoking = 0 if heske6==1 & heska==2 & wave==5
2010  * Assign the number 1 if the participant reported that they were no longer smoking cigarettes by
      Wave 5 and reported that they do not smoke cigarettes at all nowadays at Wave 5
2011  replace smoking = 1 if heske6==2 & heska==2 & wave==5
2012
2013  * Assign the number 1 if the participant reported that they stopped smoking between Wave 5 and Wave
      6 and reported that they do not smoke cigarettes at all nowadays at Wave 6
2014  replace smoking = 1 if heske==3 & heska==2 & wave==6
2015  * Assign the number 0 if the participant reported never having smoked cigarettes and reported that
```

```
they do not smoke cigarettes at all nowadays at Wave 6
replace smoking = 0 if heske7==1 & heska==2 & wave==6
* Assign the number 1 if the participant reported that they were no longer smoking cigarettes by Wave 6 and reported that they do not smoke cigarettes at all nowadays at Wave 6
replace smoking = 1 if heske7==2 & heska==2 & wave==6

* Assign the number 1 if the participant reported that they stopped smoking between Wave 6 and Wave 7 and reported that they do not smoke cigarettes at all nowadays at Wave 7
replace smoking = 1 if heske==3 & heska==2 & wave==7
* Assign the number 0 if the participant reported never having smoked cigarettes and reported that they do not smoke cigarettes at all nowadays at Wave 7
replace smoking = 0 if heske8==1 & heska==2 & wave==7
* Assign the number 1 if the participant reported that they were no longer smoking cigarettes by Wave 7 and reported that they do not smoke cigarettes at all nowadays at Wave 7
replace smoking = 1 if heske8==2 & heska==2 & wave==7

* Assign the number 1 if the participant reported that they stopped smoking between Wave 7 and Wave 8 and reported that they do not smoke cigarettes at all nowadays at Wave 8
replace smoking = 1 if heske==3 & heska==2 & wave==8
* Assign the number 0 if the participant reported never having smoked cigarettes and reported that they do not smoke cigarettes at all nowadays at Wave 8
replace smoking = 0 if heske9==1 & heska==2 & wave==8
* Assign the number 1 if the participant reported that they were no longer smoking cigarettes by Wave 8 and reported that they do not smoke cigarettes at all nowadays at Wave 8
replace smoking = 1 if heske9==2 & heska==2 & wave==8

* Assign the number 1 if the participant reported that they stopped smoking between Wave 8 and Wave 9 and reported that they do not smoke cigarettes at all nowadays at Wave 9
replace smoking = 1 if heske==3 & heska==2 & wave==9

* Assign the number 1 if the participant reported that they stopped smoking cigarettes
replace smoking = 1 if heskf==2
* Assign the number 2 if the participant reported smoking cigarettes nowadays
replace smoking = 2 if heskf==1

* Count total number of participants and observations
unique idauniq
* 15,022 individuals, 120,176 observations

* Save dataset with a new name
save datavariables04.dta

* [b45-b48] Cognitive function
* WAVE 2, 3, 4, 5, 6, 7, 8, 9
* [b45] Computed score from date questions (orientation in time)
* Generate a new variable duplicating the cfdatd variable
gen daymonth = cfdatd
* Replace variable as missing for any missing cases (coded as negative numbers in the ELSA dataset)
replace daymonth = . if daymonth<0
* Assign the number 0 if the participant answered incorrectly or didn't know the answer
replace daymonth = 0 if daymonth==2
* Generate a new variable duplicating the cfday variable
gen day = cfday
* Replace variable as missing for any missing cases (coded as negative numbers in the ELSA dataset)
replace day = . if day<0
* Assign the number 0 if the participant answered incorrectly or didn't know the answer
replace day = 0 if day==2
* Generate a new variable duplicating the cfdaty variable
gen year = cfdaty
* Replace variable as missing for any missing cases (coded as negative numbers in the ELSA dataset)
replace year = . if year<0
* Assign the number 0 if the participant answered incorrectly or didn't know the answer
replace year = 0 if year==2
```

```
2070  * Generate a new variable duplicating the cfdatm variable
2071  gen month = cfdatm
2072  * Replace variable as missing for any missing cases (coded as negative numbers in the ELSA dataset)
2073  replace month = . if month<0
2074  * Assign the number 0 if the participant answered incorrectly or didn't know the answer
2075  replace month = 0 if month==2
2076  * Generate a new variable equal to the sum of the four orientation in time items to create a total
      score
2077  gen orientation = daymonth + day + year + month
2078  gen orientation2 = cfdscr if cfdscr>=0
2079  * Generate a new variable and assign the number 0 for participants who scored 0, 1, 2, or 3 points
      on the time orientation test
2080  gen oribi = 0 if inlist(orientation,0,1,2,3)
2081  * Assign the number 1 for participants who answered all questions correctly (i.e., scored 4) on the
      time orientation test
2082  replace oribi = 1 if orientation==4
2083
2084  * WAVE 2, 3, 4, 5, 6, 7, 8, 9
2085  * [b46-b47] Word-list learning (verbal learning and recall)
2086  tab cftest
2087  tab cfwrds
2088  * Generate a new variable duplicating the cflisen variable for participants with a score from 0 to 10
2089  gen learning = cflisen if cflisen>=0
2090  * Generate a new variable duplicating the cflisd variable for participants with a score from 0 to 10
2091  gen recall = cflisd if cflisd>=0
2092
2093  * [b46] Number of words recalled immediately
2094  sum learning
2095  * Assign the number 0 for participants with scores >1 standard deviation below the mean
2096  replace learning = 0 if learning >= 0 & learning < 3.990809
2097  * Assign the number 1 for participants with scores ±1 standard deviation around the mean
2098  replace learning = 1 if learning >= 3.990809 & learning <= 7.640907
2099  * Assign the number 2 for participants with scores >1 standard deviation above the mean
2100  replace learning = 2 if learning > 7.640907 & learning != .
2101
2102  * [b47] Number of words recalled after delay
2103  sum recall
2104  * Assign the number 0 for participants with scores >1 standard deviation below the mean
2105  replace recall = 0 if recall >= 0 & recall < 2.353383
2106  * Assign the number 1 for participants with scores ±1 standard deviation around the mean
2107  replace recall = 1 if recall >= 2.353383 & recall <= 6.666215
2108  * Assign the number 2 for participants with scores >1 standard deviation above the mean
2109  replace recall = 2 if recall > 6.666215 & recall != .
2110
2111  * WAVE 2, 3, 4, 5, 7, 8, 9
2112  * [b48] Number of animals mentioned (verbal fluency)
2113  tab cfani
2114  * Generate a new variable duplicating the cfani variable for participants with a score of 0 or more
2115  gen fluency = cfani if cfani>=0
2116  sum fluency
2117  * Assign the number 0 for participants with scores >1 standard deviation below the mean
2118  replace fluency = 0 if fluency >= 0 & fluency < 13.60175
2119  * Assign the number 1 for participants with scores ±1 standard deviation around the mean
2120  replace fluency = 1 if fluency >= 13.60175 & fluency <= 27.79781
2121  * Assign the number 2 for participants with scores >1 standard deviation above the mean
2122  replace fluency = 2 if fluency > 27.79781 & fluency != .
2123
2124  * WAVE 2, 4, 6, 8, 9
2125  * [b49] Grip strength
2126  * Replace variable as missing for any missing cases (coded as negative numbers or 99 in the ELSA
      dataset)
2127  replace mmgsd1 = . if mmgsd1 < 0
2128  replace mmgsd2 = . if mmgsd2 < 0
```

```
replace mmgsd3 = . if mmgsd3 < 0
replace mmgsn1 = . if mmgsn1 < 0
replace mmgsn2 = . if mmgsn2 < 0
replace mmgsn3 = . if mmgsn3 < 0
replace mmgsd1 = . if mmgsd1==99
replace mmgsd2 = . if mmgsd2==99
replace mmgsd3 = . if mmgsd3==99
replace mmgsn1 = . if mmgsn1==99
replace mmgsn2 = . if mmgsn2==99
replace mmgsn3 = . if mmgsn3==99

* Generate a new variable equal to the maximum grip strength across all available measures
egen maxgrip = rowmax(mmgsd1 mmgsd2 mmgsd3 mmgsn1 mmgsn2 mmgsn3)

sum maxgrip
* Assign the number 0 for participants with scores >1 standard deviation below the mean
replace maxgrip = 0 if maxgrip >= 0 & maxgrip < 19.61438
* Assign the number 1 for participants with scores ±1 standard deviation around the mean
replace maxgrip = 1 if maxgrip >= 19.61438 & maxgrip <= 42.36316
* Assign the number 2 for participants with scores >1 standard deviation above the mean
replace maxgrip = 2 if maxgrip > 42.36316 & maxgrip != .

* Save dataset with a new name
save datastop.dta

* [b50-b51] (+ 4 omitted) Biomarkers
* WAVE 2, 4, 6, 8, 9
* [b50] Blood fibrinogen level (g/L)
tab cfib if cfib<0
* Generate a new variable duplicating the cfib variable
gen fibrinogen = cfib
* Replace variable as missing for any missing cases (coded as negative numbers or 99 in the ELSA
dataset)
replace fibrinogen = . if fibrinogen < 0
replace fibrinogen = . if fibrinogen > 9000
sum fibrinogen
* Assign the number 1 for participants with a fibrinogen level ≤4 g/L
replace fibrinogen = 1 if fibrinogen <= 4
* Assign the number 0 for participants with a fibrinogen level >4 g/L
replace fibrinogen = 0 if fibrinogen > 4 & fibrinogen != .

* (1 omitted) Blood HDL level (mmol/L)
tab hdl if hdl<0
* Generate a new variable duplicating the hdl variable
gen highdensity = hdl
* Replace variable as missing for any missing cases (coded as negative numbers or 99 in the ELSA
dataset)
replace highdensity = . if highdensity < 0
replace highdensity = . if highdensity > 9000
sum highdensity
* Assign the number 0 for participants with a HDL level <1 mmol/L
replace highdensity = 0 if highdensity <1
* Assign the number 1 for participants with a HDL level ≥1 mmol/L
replace highdensity = 1 if highdensity >= 1 & highdensity != .

* (1 omitted) Blood triglyceride level (mmol/L)
tab trig if trig<0
* Generate a new variable duplicating the trig variable
gen triglyceride = trig
* Replace variable as missing for any missing cases (coded as negative numbers or 99 in the ELSA
dataset)
replace triglyceride = . if triglyceride < 0
replace triglyceride = . if triglyceride > 9000
```

```
sum triglyceride
* Assign the number 1 for participants with a triglyceride level ≤2 mmol/L
replace triglyceride = 1 if triglyceride <= 2
* Assign the number 0 for participants with a triglyceride level >2 mmol/L
replace triglyceride = 0 if triglyceride > 2 & triglyceride != .

* (1 omitted) Blood LDL level (mmol/L)
tab ldl if ldl<0
* Generate a new variable duplicating the ldl variable
gen lowdensity = ldl
* Replace variable as missing for any missing cases (coded as negative numbers or 99 in the ELSA dataset)
replace lowdensity = . if lowdensity < 0
replace lowdensity = . if lowdensity > 9000
sum lowdensity
* Assign the number 1 for participants with a LDL level ≤4 mmol/L
replace lowdensity = 1 if lowdensity <= 4
* Assign the number 0 for participants with a LDL level >4 mmol/L
replace lowdensity = 0 if lowdensity > 4 & lowdensity != .

* [b51] Blood CRP level (mg/L)
tab hscrp if hscrp<0
* Generate a new variable duplicating the hscrp variable
gen CRP = hscrp
* Replace variable as missing for any missing cases (coded as negative numbers or 99 in the ELSA dataset)
replace CRP = . if CRP < 0
replace CRP = . if CRP > 9000
sum CRP
* Replace variable as missing for CRP values >20 mg/L
replace CRP = . if CRP > 20
* Assign the number 1 for participants with a CRP level ≤3 mg/L
replace CRP = 1 if CRP <= 3
* Assign the number 0 for participants with a CRP level >3 mg/L
replace CRP = 0 if CRP > 3 & CRP != .

* (1 omitted) Blood glycated haemoglobin level (%)
* Wave 2, 4 = %, Wave 6, 8, 9 = mmol/mol
tab hba1c if hba1c<0
* Generate a new variable duplicating the hba1c variable
gen glycated = hba1c
* Replace variable as missing for any missing cases (coded as negative numbers or 99 in the ELSA dataset)
replace glycated = . if glycated < 0
replace glycated = . if glycated > 9000
sum glycated if wave==2
sum glycated if wave==4
sum glycated if wave==6
sum glycated if wave==8
sum glycated if wave==9
* Transform HbA1c levels (mmol/mol) in Wave 6, 8, and 9 into % to match data in Wave 2 and 4
replace glycated = ((glycated/10.929)+2.15) if inlist(wave,6,8,9)
sum glycated if wave==2
sum glycated if wave==4
sum glycated if wave==6
sum glycated if wave==8
sum glycated if wave==9
* Assign the number 1 for participants with a HbA1c level <6.5 %
replace glycated = 1 if glycated < 6.5
* Assign the number 0 for participants with a HbA1c level ≥6.5 %
replace glycated = 0 if glycated >= 6.5 & glycated != .

* Save dataset with a new name
```

```
save alldataefa.dta

******************
***DATA ANALYSIS***
******************

* Keep variables required for analyses
keep idauniq wave revscacta revscactb revscactc revscactd scpt04 scpt05 headldr headlwa headlba
headlea headlbe headlwc headlma headlpr headlsh headlph headlme headlho headlmo hemobwa hemobsi
hemobch hemobcs hemobcl hemobst hemobre hemobpu hemobli hemobpi limiting scorg01 scorg02 scorg03
scorg04 scorg05 scorg06 scorg07 scorg08 balance repcstest gaittest revhehelf psceda pscedb pscedc
pscedd pscede pscedf pscedg pscedh revscfeela revscfeelb revscfeelc oribi learning recall fluency
maxgrip fibrinogen highdensity triglyceride lowdensity CRP glycated
* Rename variables to more convenient forms
rename revscacta scactarev
rename revscactb scactbrev
rename revscactc scactcrev
rename revscactd scactdrev
rename revhehelf hehelfrev
rename revscfeela scfeelarev
rename revscfeelb scfeelbrev
rename revscfeelc scfeelcrev
* Save dataset with a new name
save efanew.dta
* Keep data from Wave 2 only
keep if wave==2
* Drop unnecessary variable
drop wave
* Count total number of participants
unique idauniq
* 15,022 individuals
* Save dataset with a new name
save baseline.dta
* Generate a new variable equal to the sum of variables with available "non-missing" data for each
participant
egen nmcount = rownonmiss(_all), strok
tab nmcount
* Drop observation if > 25% missing values across the 64 variables of interest (idauniq and wave
data are complete for all observations)
drop if nmcount<49
* Drop unnecessary variable
drop nmcount
* Count total number of participants
unique idauniq
* 7,660 individuals
* Generate a new variable equal to the sum of variables with missing data for each participant
egen nmcount = rowmiss(_all)
tab nmcount
* Drop unnecessary variable
drop nmcount
* Set the seed
set seed 1234
* Generate random numbers
gen random = uniform()
sort random
* Assign 30% of the total sample to the validation sub-sample
gen byte validation = _n <= 2298
* Save dataset with a new name
save efatotal.dta
* Keep participants assigned to the developmental sub-sample
keep if validation == 0
* Count total number of participants
unique idauniq
```

```
* 5,362 individuals
* Save developmental dataset
save developmentalnew.dta
* Use efatotal.dta dataset
use efatotal.dta
* Keep participants assigned to the validation sub-sample
keep if validation == 1
* Count total number of participants
unique idauniq
* 2,298 individuals
* Save validation dataset
save validationnew.dta

* Use developmental dataset
use developmentalnew.dta
* Convert Stata data into a data file and Mplus input file
stata2mplus using developmentalfinal.dta

* Use validation dataset
use validationnew.dta
* Convert Stata data into a data file and Mplus input file
stata2mplus using validationfinal.dta

* Use efanew.dta dataset
use efanew.dta
* Drop variables omitted from final metric
drop pscedb pscedh headlmo scfeelarev scfeelbrev scfeelcrev balance headlme triglyceride lowdensity highdensity glycated scorg06
* Save dataset with a new name
save MLIRT.dta
* Count total number of participants and observations
unique idauniq
* 15,022 individuals, 120,176 observations
* Generate a new variable equal to the sum of variables with available "non-missing" data for each participant
egen nmcount = rownonmiss(_all), strok
tab nmcount
* Drop observation if > 50% missing values across the 51 variables of interest (idauniq and wave data are complete for all observations)
keep if nmcount>=27.5
* Drop unnecessary variable
drop nmcount
* Count total number of participants and observations
unique idauniq
* 14,755 individuals, 66,133 observations
* Generate a new variable equal to the sum of variables with missing data for each participant
egen nmcount = rowmiss(_all)
tab nmcount
* Drop unnecessary variable
drop nmcount
* Save dataset with a new name
save MLIRTtouse.dta
* Export Stata data to .csv file
export delimited using "", nolabel replace

* Use alldataefa.dta dataset
use alldataefa.dta

* WAVE 2, 3, 4, 5, 6, 8, 9
* Socio-economic covariate - Occupational class
* Excluded Never worked and long-term unemployed
* Replace variable as missing for any missing cases (coded as negative numbers or 99 in the ELSA dataset)
```

```
2364  replace nssec8 = . if nssec8<0
2365  replace nssec8 = . if nssec8 == 99
2366  * Generate a new variable
2367  gen mynssec3 = .
2368  * Assign the number 2 if the participant's current or most recent occupation was coded as: Higher
      managerial, administrative and professional occupations; or Lower managerial, administrative and
      professional occupations
2369  replace mynssec3 = 2 if inlist(nssec8,1,2)
2370  * Assign the number 1 if the participant's current or most recent occupation was coded as:
      Intermediate occupation; or Small employers and own account workers
2371  replace mynssec3 = 1 if inlist(nssec8,3,4)
2372  * Assign the number 0 if the participant's current or most recent occupation was coded as: Lower
      supervisory and technical occupations; or Semi-routine occupations; or Routine occupations
2373  replace mynssec3 = 0 if inlist(nssec8,5,6,7)
2374  * Overwrite dataset, by replacing the previously saved file
2375  save alldataefa.dta, replace
2376
2377  * Keep variables required for analyses
2378  keep idauniq wave indager fqethnr education totwq5_bu_s disex scako smoking activity2 mynssec3 QoL
2379  * Save dataset with a new name
2380  save MLIRTGMM.dta
2381  * One-to-one merge of data in memory with MLIRTdata.dta (exported from RStudio) on participant ID
2382  merge 1:1 idauniq wave using MLIRTdata.dta
2383  * Sort from lowest to highest participant ID
2384  sort idauniq
2385  * Save dataset with a new name
2386  save MLIRTdatafull.dta
2387  tab AHA
2388  * Drop observations with missing AHA scores
2389  drop if AHA==.
2390  * Count total number of participants and observations
2391  unique idauniq
2392  * 14,755 individuals, 66,133 observations
2393  * Save dataset with a new name
2394  save MLIRTfullAHA.dta
2395  sum indager, d
2396  sum indager if wave==2, d
2397  * Generate a variable that assigns the observation number (i.e., 1 for first data collection
      timepoint, 2 for second data collection timepoint etc.) to each row by participant ID
2398  bysort idauniq (wave): gen obsnr = _n
2399  * Generate a variable that assigns the number of total observations to each row of data for a given
      participant
2400  bysort idauniq: gen obscount = _N
2401  tab obscount
2402  * Include participants with AHA scores in at least two waves
2403  keep if obscount!=1
2404  * Count total number of participants and observations
2405  unique idauniq
2406  * 11,739 individuals, 63,117 observations
2407  * Keep necessary variables
2408  keep idauniq wave indager fqethnr education totwq5_bu_s disex scako smoking activity2 mynssec3 QoL
      AHA obsnr obscount
2409  sum idauniq if obsnr==1 & wave==2
2410  * Generate a variable that assigns the number 1 to the row representing participants' first
      observation if this corresponds to Wave 2 (baseline)
2411  bysort idauniq (wave): gen firstwave = 1 if obsnr==1 & wave==2
2412  * Carry the value of this last variable forwards to the remainder of a participant's observations
2413  bysort idauniq: gen firstwave_cons = firstwave[1]
2414  * Include participants with AHA scores in at least two waves, including at baseline
2415  keep if firstwave_cons==1
2416  * Count total number of participants and observations
2417  unique idauniq
2418  * 7,416 individuals, 42,716 observations
```

```
2419
2420  * Generate a new variable duplicating the biological sex variable at Wave 2
2421  gen sex = disex if wave==2
2422  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
2423  tsset idauniq wave
2424  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 2 to the
      follow-up waves available for that participant) by participant ID
2425  bysort idauniq: carryforward sex, replace
2426
2427  * Generate a new variable duplicating the education variable at Wave 2
2428  gen qualifications = education if wave==2
2429  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
2430  tsset idauniq wave
2431  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 2 to the
      follow-up waves available for that participant) by participant ID
2432  bysort idauniq: carryforward qualifications, replace
2433
2434  * Generate a new variable duplicating the wealth variable at Wave 2
2435  gen wealth = totwq5_bu_s if wave==2
2436  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
2437  tsset idauniq wave
2438  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 2 to the
      follow-up waves available for that participant) by participant ID
2439  bysort idauniq: carryforward wealth, replace
2440
2441  * Generate a new variable duplicating the age variable at Wave 2
2442  gen age = indager if wave==2
2443  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
2444  tsset idauniq wave
2445  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 2 to the
      follow-up waves available for that participant) by participant ID
2446  bysort idauniq: carryforward age, replace
2447
2448  * Generate a new variable duplicating the ethnicity variable at Wave 2
2449  gen ethnicity = fqethnr if wave==2
2450  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
2451  tsset idauniq wave
2452  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 2 to the
      follow-up waves available for that participant) by participant ID
2453  bysort idauniq: carryforward ethnicity, replace
2454
2455  * Generate a new variable duplicating the occupational class variable at Wave 2
2456  gen nssec3 = mynssec3 if wave==2
2457  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
2458  tsset idauniq wave
2459  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 2 to the
      follow-up waves available for that participant) by participant ID
2460  bysort idauniq: carryforward nssec3, replace
2461
2462  * Save dataset with a new name
2463  save GMMcovdataSES.dta
2464  * Keep variables required for analyses
2465  keep idauniq wave sex qualifications wealth age ethnicity nssec3 QoL AHA
2466  * Reshape data into wide format for observations identified by participant ID and add wave as an
      identifying time period
2467  reshape wide QoL AHA, j(wave) i(idauniq)
2468  * Save dataset with a new name
2469  save GMMcovdatawideSES.dta
2470  sum QoL9, d
2471  * Generate a new variable and assign the number 0 if a participant's quality-of-life score is below
      the sample median
2472  gen QoL9binary = 0 if QoL9<43
2473  * Assign the number 1 if a participant's quality-of-life score is above or equal to the sample median
```

```
2474  replace QoL9binary = 1 if QoL9 >=43 & QoL9!=.
2475  * Save dataset with a new name
2476  save GMMcovbinarySES.dta
2477
2478  * Dummy variables for conditional GMM
2479  * Education
2480  * Medium education (i.e., school qualifications) (coded as 1) versus low (i.e., no formal
      qualifications) or high (i.e., higher education) education (coded as 0)
2481  gen mediumed = 0 if inlist(qualifications,0,2)
2482  replace mediumed = 1 if qualifications == 1
2483  * High education (coded as 1) versus low or medium education (coded as 0)
2484  gen highed = 0 if inlist(qualifications,0,1)
2485  replace highed = 1 if qualifications == 2
2486  * Occupational class
2487  * Intermediate occupations (coded as 1) versus lower or higher occupations (coded as 0)
2488  gen mediumocc = 0 if inlist(nssec3,0,2)
2489  replace mediumocc = 1 if nssec3 == 1
2490  * Higher occupations (coded as 1) versus lower or intermediate occupations (coded as 0)
2491  gen highocc = 0 if inlist(nssec3,0,1)
2492  replace highocc = 1 if nssec3 == 2
2493  * Wealth
2494  * 2nd quintile (coded as 1) versus 1st, 3rd, 4th, or 5th quintile (coded as 0)
2495  gen quint2 = 0 if inlist(wealth,1,3,4,5)
2496  replace quint2 = 1 if wealth == 2
2497  * 3rd quintile (coded as 1) versus 1st, 2nd, 4th, or 5th quintile (coded as 0)
2498  gen quint3 = 0 if inlist(wealth,1,2,4,5)
2499  replace quint3 = 1 if wealth == 3
2500  * 4th quintile (coded as 1) versus 1st, 2nd, 3rd, or 5th quintile (coded as 0)
2501  gen quint4 = 0 if inlist(wealth,1,2,3,5)
2502  replace quint4 = 1 if wealth == 4
2503  * 5th quintile (coded as 1) versus 1st, 2nd, 3rd, or 4th quintile (coded as 0)
2504  gen quint5 = 0 if inlist(wealth,1,2,3,4)
2505  replace quint5 = 1 if wealth == 5
2506  * Save dataset with a new name
2507  save GMMcovbinarySESdu.dta
2508  * Convert Stata data into a data file and Mplus input file
2509  stata2mplus using GMMcovbinarySESdu.dta
2510
2511  * Use MLIRTfullAHA.dta dataset
2512  use MLIRTfullAHA.dta
2513  * Count total number of participants and observations
2514  unique idauniq
2515  * 14,755 individuals, 66,133 observations
2516  * Generate a variable that assigns the observation number (i.e., 1 for first data collection
      timepoint, 2 for second data collection timepoint etc.) to each row by participant ID
2517  bysort idauniq (wave): gen obsnr = _n
2518  * Generate a new variable duplicating the biological sex variable at Wave 2
2519  gen sex = disex if obsnr==1
2520  * Declare a panel dataset with participant ID "idauniq" and time variable "wave"
2521  tsset idauniq wave
2522  * Carryforward observations with respect to the time variable "wave" (i.e., from Wave 2 to the
      follow-up waves available for that participant) by participant ID
2523  bysort idauniq: carryforward sex, replace
2524  * Keep variables required for Receiver Operating Characteristic (ROC) analyses
2525  keep idauniq wave sex QoL AHA
2526  * Save dataset with a new name
2527  save ROCdata.dta
2528  sum QoL, d
2529  * Generate a new variable and assign the number 0 if a participant's quality-of-life score is below
      the sample median
2530  gen QoLbinary = 0 if QoL < 43
2531  * Assign the number 1 if a participant's quality-of-life score is above or equal to the sample median
2532  replace QoLbinary = 1 if QoL >= 43 & QoL !=.
```

```
* Overwrite dataset, by replacing the previously saved file
save ROCdata.dta, replace
* Drop unnecessary variable
drop QoL
* Reshape data into wide format for observations identified by participant ID and add wave as an identifying time period
reshape wide AHA QoLbinary, j(wave) i (idauniq)
* Save dataset with a new name
save ROC92wide.dta

* ROC analyses, clustering at the participant level and adjusting the control distribution for biological sex
* Quality-of-life at Wave 3
rocreg QoLbinary3 AHA2, probit ml ctrlcov(sex) ctrlmodel(linear) cluster(idauniq)
* Quality-of-life at Wave 6
rocreg QoLbinary6 AHA2, probit ml ctrlcov(sex) ctrlmodel(linear) cluster(idauniq)
* Quality-of-life at Wave 9
rocreg QoLbinary9 AHA2, probit ml ctrlcov(sex) ctrlmodel(linear) cluster(idauniq)

* Use MLIRTfullAHA.dta dataset
use MLIRTfullAHA.dta
* Two-level mixed-effects linear regression of AHA scores on lifestyle behaviours (adjusted for covariates), with random intercepts by participant ID
mixed AHA indager i.disex i.fqethnr i.scako i.smoking i.activity2 || idauniq:
* Store estimates for later use
estimates store randint
* Fit indices
estimates stats
* Two-level mixed-effects linear regression of AHA scores on lifestyle behaviours (adjusted for covariates), with random intercepts by participant ID and a random slope according to wave
mixed AHA indager i.disex i.fqethnr i.scako i.smoking i.activity2 || idauniq: wave
* Store estimates for later use
estimates store randslope
* Fit indices
estimates stats
* Likelihood ratio test
lrtest randslope randint

* Use MLIRTfullAHA.dta dataset
use MLIRTfullAHA.dta
* Count total number of participants and observations
unique idauniq
* 14,755 individuals, 66,133 observations
* Generate a variable that assigns the observation number (i.e., 1 for first data collection timepoint, 2 for second data collection timepoint etc.) to each row by participant ID
bysort idauniq (wave): gen obsnr = _n
* Generate a variable that assigns the number 1 to the row representing participants' first observation if this corresponds to Wave 2 (baseline)
bysort idauniq (wave): gen firstwave = 1 if obsnr==1 & wave==2
* Carry the value of this last variable forwards to the remainder of a participant's observations
bysort idauniq: gen firstwave_cons = firstwave[1]
* Include participants with at least some data at baseline
keep if firstwave_cons==1
* Generate a variable that assigns the number of total observations to each row of data for a given participant
bysort idauniq: gen obscount = _N
tab pycount
* Count total number of participants and observations
unique idauniq
* 8,670 individuals, 43,970 observations
* Keep necessary variables
keep idauniq wave indager fqethnr education totwq5_bu_s disex scako smoking activity2 mynssec3 QoL AHA obsnr obscount firstwave firstwave_cons
```

```
* Generate a new variable and assign the number 0 to participants with data at a single wave
gen total = 0 if obscount == 1
* Assign the number 1 to participants with AHA scores in at least two waves, including baseline
replace total = 1 if obscount >1
* Keep data from Wave 2 only
keep if wave==2
* Count total number of participants
unique idauniq
* 8,670 individuals
tab total
* Baseline characteristics of the complete-case sample used for the GMM analyses versus the sample excluded due to missing data (unweighted)
sum indager if total == 0
sum indager if total == 1
tab fqethnr if total == 0
tab fqethnr if total == 1
tab education if total == 0
tab education if total == 1
tab totwq5_bu_s if total == 0
tab totwq5_bu_s if total == 1
tab disex if total == 0
tab disex if total == 1
tab scako if total == 0
tab scako if total == 1
tab smoking if total == 0
tab smoking if total == 1
tab activity2 if total == 0
tab activity2 if total == 1
tab mynssec3 if total == 0
tab mynssec3 if total == 1
ttest indager, by(total)
tabulate fqethnr total, chi2
tabulate education total, chi2
tabulate totwq5_bu_s total, chi2
tabulate disex total, chi2
tabulate scako total, chi2
tabulate smoking total, chi2
tabulate activity2 total, chi2
tabulate mynssec3 total, chi2

* Use GMMcovdataSES.dta dataset
use GMMcovdataSES.dta
* Drop observations with missing data on socio-economic or demographic covariates
drop if sex==. | qualifications==. | wealth==. | age==. | ethnicity==. | nssec3 ==.
* Count total number of participants and observations
unique idauniq
* 6,539 individuals, 37,768 observations
* Drop unnecessary variables
drop obsnr obscount
* Generate a variable that assigns the observation number (i.e., 1 for first data collection timepoint, 2 for second data collection timepoint etc.) to each row by participant ID
bysort idauniq (wave): gen obsnr = _n
* Generate a variable that assigns the number of total observations to each row of data for a given participant
bysort idauniq: gen obscount = _N
tab wave
* Summarise number of observations per participant
tab obscount if wave==2
sum obscount if wave==2
* Keep necessary variables
keep idauniq wave sex qualifications wealth age ethnicity nssec3 AHA QoL scako smoking activity2
* Reshape data into wide format for observations identified by participant ID and add wave as an identifying time period
```

```
reshape wide QoL AHA scako smoking activity2, j(wave) i(idauniq)
* Save dataset with a new name
save descbyclassGMMcovdataSESwide.dta

* Import posterior probabilities of class membership following the conditional 3-class GMM
import excel "", sheet("") firstrow
* Save dataset with a new name
save classesfinal.dta
* Use descbyclassGMMcovdataSESwide.dta dataset
use descbyclassGMMcovdataSESwide.dta
* One-to-one merge of data in memory with classesfinal.dta on participant ID
merge 1:1 idauniq using classesfinal.dta
* Save dataset with a new name
save NEWclassdescfinal.dta

* Summarise the AHA scores at each wave
sum AHA2, d
sum AHA3, d
sum AHA4, d
sum AHA5, d
sum AHA6, d
sum AHA7, d
sum AHA8, d
sum AHA9, d

* Summarise posterior probabilities for each latent class
sum posterior1 if class==1
sum posterior2 if class==2
sum posterior3 if class==3

* Descriptive statistics stratified by class membership
sum age if class==1
sum age if class==2
sum age if class==3

tab sex if class==1
tab sex if class==2
tab sex if class==3

tab ethnicity if class==1
tab ethnicity if class==2
tab ethnicity if class==3

tab wealth if class==1
tab wealth if class==2
tab wealth if class==3

tab qualifications if class==1
tab qualifications if class==2
tab qualifications if class==3

tab nssec3 if class==1
tab nssec3 if class==2
tab nssec3 if class==3
```