```
*Computer code used for the PLoS ONE paper:
*Piirtola et al (2018), Association of current and former smoking with body mass index:
*a study of smoking discordant twin pairs from 21 twin cohorts Association

**Figure 1

*Creating the basic data
use "CODATWINS DATA SMOKING FOR 21 COHORTS", clear
drop if age>69 // keeping only those >= 70-years of age
sort gindid myeargroup age // keeping only one BMI/ smoking measurement per each 10-year measurement period
quietly by gindid myeargroup: gen dup= cond(_N==1, 0,_n)
tab dup
drop if dup>1
drop dup

** -> n=156,593 twin individuals (51% male) with m=230,378 BMI and smoking measurements

*Creating data for within pair analyses
drop if zygosdz==3 // opposite-sex pairs are removed

sort gtwinid myeargroup // UNMATCHED PAIRS WITHIN myeargroup ARE REMOVED
quietly by gtwinid myeargroup: gen dup = cond(_N==1,0,_n)
tab dup
drop if dup==0
drop dup
sum bmi

*Illogical zygosity is droped
misstable sum zygosdz
list gindid gtwinid cohort zyg sex zygosdz myeargroup if zygosdz==.
drop if zygosdz==.
sort gtwinid myeargroup // UNMATCHED PAIRS WITHIN myeargroup ARE REMOVED AGAIN
quietly by gtwinid myeargroup: gen dup = cond(_N==1,0,_n)
tab dup
drop if dup==0
drop dup
sum bmi

*Number of pairs in the long data format
sort gtwinid
quietly by gtwinid: gen dup = cond(_N==1,0,_n)
tab dup sex, row
tab dup zyg, row
by sex zyg, sort: tab dup // number of full pairs is dup==1
drop dup

*-> m=80,718 pairwise BMI and smoking measurements in a total of 55,296 twin pairs (47% MZ pairs)

save "CODATwins data smoking_BMI"

use "CODATwins data smoking_BMI", clear

*Number of repeated measurement in the data
sort gindid
quietly by gindid: gen dup= cond(_N==1, 0,_n)
tab dup
by sex, sort: tab dup
drop dup

*Interaction effect between sex and smoking
```

```
regress bmi i.smoke i.sex i.smoke#i.sex age, cl(gtwinid)
est store A
regress bmi i.smoke i.sex age, cl(gtwinid)
est store B
lrtest A B, force

*Interaction effect between sex and 10-year time period
regress bmi i.smoke i.myeargroup i.sex i.cohort i.smoke#i.myeargroup age, cl(gtwinid)
est store C
regress bmi i.smoke i.myeargroup i.sex i.cohort age, cl(gtwinid)
est store D
lrtest C D, force

*Table 1

by myeargroup, sort: sum age
by myeargroup sex, sort: sum age
by area sex smoke, sort: sum bmi age
by myeargroup sex smoke, sort: sum bmi age
by sex smoke, sort: sum bmi
by sex, sort: sum age
by area sex, sort: sum age
by myeargroup sex, sort: sum age
by sex, sort: tab bmi myeargroup

* S1 Table

by smoke, sort: sum bmi
by area smoke, sort: sum bmi
by area sex smoke, sort: sum bmi
by sex smoke, sort: sum bmi
by cohort smoke, sort: sum bmi
by cohort sex smoke , sort: sum bmi

**S2-S4 Tables
*Individual level associations

regress bmi i.smoke age age2 i.cohort i.sex, cl(gtwinid)
margins, within(sex smoke)
by sex, sort: regress bmi i.smoke age age2 i.cohort, cl(gtwinid) // reference group never smokers
by sex, sort: regress bmi ib(2).smoke age age2 i.cohort, cl(gtwinid) //reference group current smokers

*BY MYEARGROUP (reference group never-smokers)

tab smoke

regress bmi i.smoke age age2 if myeargroup==1960, cl(gtwinid) // obs. only men in one cohort
margins, within(smoke)

regress bmi i.smoke age age2 i.cohort i.sex if myeargroup==1970, cl(gtwinid)
margins, within(sex smoke)

regress bmi i.smoke age age2 i.cohort i.sex if myeargroup==1980, cl(gtwinid)
margins, within(sex smoke)

regress bmi i.smoke age age2 i.cohort i.sex if myeargroup==1990, cl(gtwinid)
margins, within(sex smoke)

regress bmi i.smoke age age2 i.cohort i.sex if myeargroup==2000, cl(gtwinid)
margins, within(sex smoke)
```

```
*By sex

sort sex myeargroup
by sex myeargroup, sort: regress bmi i.smoke age age2 i.cohort, cl(gtwinid)

sort sex // all time periods
by sex, sort: regress bmi i.smoke age age2 i.cohort, cl(gtwinid)

*BY MYEARGROUP (ref group smokers)

tab smoke

regress bmi ib(2).smoke age age2 if myeargroup==1960, cl(gtwinid) // obs. only men in one cohort
margins, within(smoke)

regress bmi ib(2).smoke age age2 i.cohort i.sex if myeargroup==1970, cl(gtwinid)
margins, within(sex smoke)

regress bmi ib(2).smoke age age2 i.cohort i.sex if myeargroup==1980, cl(gtwinid)
margins, within(sex smoke)

regress bmi ib(2).smoke age age2 i.cohort i.sex if myeargroup==1990, cl(gtwinid)
margins, within(sex smoke)

regress bmi ib(2).smoke age age2 i.cohort i.sex if myeargroup==2000, cl(gtwinid)
margins, within(sex smoke)

*By sex
sort sex myeargroup
by sex myeargroup, sort: regress bmi ib(2).smoke age age2 i.cohort, cl(gtwinid)

sort sex  // over years
by sex, sort: regress bmi ib(2).smoke age age2 i.cohort, cl(gtwinid)

**Within-pair associations

**Number of paired tests in the data
sort cohort twinid myeargroup
quietly by cohort twinid myeargroup: gen dup = cond(_N==1,0,_n)
tab dup
drop dup

*Reference group never-smokers
xtset gtwinid
by sex, sort: xtreg bmi i.smoke age age2, cl (gtwinid) fe // all men and women in all periods, cl (gtwinid) for repeated measuremnts in some pairs
by sex zyg, sort: xtreg bmi i.smoke age age2, cl(gtwinid) fe // zyg by sex in all periods, cl (gtwinid) for repeated measuremnts in some pairs
by myeargroup sex, sort: xtreg bmi i.smoke age age2, fe // by sex all 10-year periods (all pairs)
by myeargroup sex zyg, sort: xtreg bmi i.smoke age age2, fe // by all different myeargroups, sex and zyg

*reference group smokers
xtset gtwinid
by sex, sort: xtreg bmi ib(2).smoke age age2, cl (gtwinid) fe // all men and women in all periods, cl (gtwinid) for repeated measuremnts in some pairs
by sex zyg, sort: xtreg bmi ib(2).smoke age age2, cl(gtwinid) fe // zyg by sex in all periods, cl (gtwinid) for repeated measurements in some pairs
by myeargroup sex, sort: xtreg bmi ib(2).smoke age age2, fe // by sex all 10-year periods (all pairs)
by myeargroup sex zyg, sort: xtreg bmi ib(2).smoke age age2, fe // by all different myeargroups, sex and zyg
```

```
***Number of discordant pairs in the analyses (Table 1, S2-S4 Tables)

sort gtwinid myeargroup smoke
quietly by gtwinid myeargroup: gen dup = cond(_N==1,0,_n)
tab dup

*Reshaping data to wide format

keep gtwinid cohort gindid twinid zyg bmi sex myeargroup smoke age dup
reshape wide age gindid twinid bmi smoke , i(gtwinid myeargroup) j(dup)

*Removing smoking concordant twin pairs per 10-year period

drop if smoke1==smoke2&myeargroup==2000
drop if smoke1==smoke2&myeargroup==1990
drop if smoke1==smoke2&myeargroup==1980
drop if smoke1==smoke2&myeargroup==1970
drop if smoke1==smoke2&myeargroup==1960

tab zyg sex

*Now only smoking discordant twin pairs are in the data

tab zyg if myeargroup==2000
tab smoke1 smoke2 if myeargroup==2000
tab smoke1 smoke2 if myeargroup==2000&zyg==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==2000&zyg==2 // DZ pairs

tab zyg if myeargroup==1990
tab smoke1 smoke2 if myeargroup==1990
tab smoke1 smoke2 if myeargroup==1990&zyg==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==1990&zyg==2 // DZ pairs

tab zyg if myeargroup==1980
tab smoke1 smoke2 if myeargroup==1980
tab smoke1 smoke2 if myeargroup==1980&zyg==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==1980&zyg==2 // DZ pairs

tab zyg if myeargroup==1970
tab smoke1 smoke2 if myeargroup==1970
tab smoke1 smoke2 if myeargroup==1970&zyg==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==1970&zyg==2 // DZ pairs

tab zyg if myeargroup==1960
tab smoke1 smoke2 if myeargroup==1960
tab smoke1 smoke2 if myeargroup==1960&zyg==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==1960&zyg==2 // DZ pairs

tab myeargroup sex

tab zyg if myeargroup==2000&sex==1
tab smoke1 smoke2 if myeargroup==2000&sex==1
tab smoke1 smoke2 if myeargroup==2000&zyg==1&sex==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==2000&zyg==2&sex==1 // DZ pairs

tab zyg if myeargroup==2000&sex==2
tab smoke1 smoke2 if myeargroup==2000&sex==2
tab smoke1 smoke2 if myeargroup==2000&zyg==1&sex==2 // MZ pairs
tab smoke1 smoke2 if myeargroup==2000&zyg==2&sex==2 // DZ pairs
```

```
tab zyg if myeargroup==1990&sex==1
tab smoke1 smoke2 if myeargroup==1990&sex==1
tab smoke1 smoke2 if myeargroup==1990&zyg==1&sex==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==1990&zyg==2&sex==1 // DZ pairs

tab zyg if myeargroup==1990&sex==2
tab smoke1 smoke2 if myeargroup==1990&sex==2
tab smoke1 smoke2 if myeargroup==1990&zyg==1&sex==2 // MZ pairs
tab smoke1 smoke2 if myeargroup==1990&zyg==2&sex==2 // DZ pairs

tab zyg if myeargroup==1980&sex==1
tab smoke1 smoke2 if myeargroup==1980&sex==1
tab smoke1 smoke2 if myeargroup==1980&zyg==1&sex==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==1980&zyg==2&sex==1 // DZ pairs

tab zyg if myeargroup==1980&sex==2
tab smoke1 smoke2 if myeargroup==1980&sex==2
tab smoke1 smoke2 if myeargroup==1980&zyg==1&sex==2 // MZ pairs
tab smoke1 smoke2 if myeargroup==1980&zyg==2&sex==2 // DZ pairs

tab zyg if myeargroup==1970&sex==1
tab smoke1 smoke2 if myeargroup==1970&sex==1
tab smoke1 smoke2 if myeargroup==1970&zyg==1&sex==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==1970&zyg==2&sex==1 // DZ pairs

tab zyg if myeargroup==1970&sex==2
tab smoke1 smoke2 if myeargroup==1970&sex==2
tab smoke1 smoke2 if myeargroup==1970&zyg==1&sex==2 // MZ pairs
tab smoke1 smoke2 if myeargroup==1970&zyg==2&sex==2 // DZ pairs

tab zyg if myeargroup==1960&sex==1
tab smoke1 smoke2 if myeargroup==1960&sex==1
tab smoke1 smoke2 if myeargroup==1960&zyg==1&sex==1 // MZ pairs
tab smoke1 smoke2 if myeargroup==1960&zyg==2&sex==1 // DZ pairs

save "Data wide smoking discordant pairs.dta"

*METAN procedure used for meta-analysis

use "Meta_data_for_stata_current_never"

destring beta, gen(beta_n)
destring lcl, gen(lcl_n)
destring hcl, gen(hcl_n)

metan beta_n lcl_n hcl_n if sex==1&group==3 // men ind
metan beta_n lcl_n hcl_n if sex==1&group==2 // men DZ
metan beta_n lcl_n hcl_n if sex==1&group==1 // men MZ

metan beta_n lcl_n hcl_n if sex==2&group==3 // women ind
metan beta_n lcl_n hcl_n if sex==2&group==2 // women DZ
metan beta_n lcl_n hcl_n if sex==2&group==1 // women MZ

metan beta_n lcl_n hcl_n if sex==1&period==1960
metan beta_n lcl_n hcl_n if sex==1&period==1970
metan beta_n lcl_n hcl_n if sex==1&period==1980
metan beta_n lcl_n hcl_n if sex==1&period==1990
metan beta_n lcl_n hcl_n if sex==1&period==2000
```

```
metan beta_n lcl_n hcl_n if sex==2&period==1970
metan beta_n lcl_n hcl_n if sex==2&period==1980
metan beta_n lcl_n hcl_n if sex==2&period==1990
metan beta_n lcl_n hcl_n if sex==2&period==2000

use "Meta_data_for_stata_former_current"

destring beta, gen(beta_n)
destring lcl, gen(lcl_n)
destring hcl, gen(hcl_n)

metan beta_n lcl_n hcl_n if sex==1&group==3 // men ind
metan beta_n lcl_n hcl_n if sex==1&group==2 // men DZ
metan beta_n lcl_n hcl_n if sex==1&group==1 // men MZ

metan beta_n lcl_n hcl_n if sex==2&group==3 // women ind
metan beta_n lcl_n hcl_n if sex==2&group==2 // women DZ
metan beta_n lcl_n hcl_n if sex==2&group==1 // women MZ

metan beta_n lcl_n hcl_n if sex==1&period==1960
metan beta_n lcl_n hcl_n if sex==1&period==1970
metan beta_n lcl_n hcl_n if sex==1&period==1980
metan beta_n lcl_n hcl_n if sex==1&period==1990
metan beta_n lcl_n hcl_n if sex==1&period==2000

metan beta_n lcl_n hcl_n if sex==2&period==1970
metan beta_n lcl_n hcl_n if sex==2&period==1980
metan beta_n lcl_n hcl_n if sex==2&period==1990
metan beta_n lcl_n hcl_n if sex==2&period==2000

use "Meta_data_for_stata_former_never"

destring beta, gen(beta_n)
destring lcl, gen(lcl_n)
destring hcl, gen(hcl_n)

metan beta_n lcl_n hcl_n if sex==1&group==3 // men ind
metan beta_n lcl_n hcl_n if sex==1&group==2 // men DZ
metan beta_n lcl_n hcl_n if sex==1&group==1 // men MZ

metan beta_n lcl_n hcl_n if sex==2&group==3 // women ind
metan beta_n lcl_n hcl_n if sex==2&group==2 // women DZ
metan beta_n lcl_n hcl_n if sex==2&group==1 // women MZ

metan beta_n lcl_n hcl_n if sex==1&period==1960
metan beta_n lcl_n hcl_n if sex==1&period==1970
metan beta_n lcl_n hcl_n if sex==1&period==1980
metan beta_n lcl_n hcl_n if sex==1&period==1990
metan beta_n lcl_n hcl_n if sex==1&period==2000

metan beta_n lcl_n hcl_n if sex==2&period==1970
metan beta_n lcl_n hcl_n if sex==2&period==1980
metan beta_n lcl_n hcl_n if sex==2&period==1990
metan beta_n lcl_n hcl_n if sex==2&period==2000
```