```
1
2
3   ****************************************
4   ****************************************
5   **#///FZ - JUN 2023 - R1 DIABETES CARE**
6   ****************************************
7   ****************************************
8
9   **#***People between 40-74 [screening diabetes NHS]****
10  cap use "...incidence_db1.dta", replace
11  egen float agec = cut(agein), at(0 40 75 150) icodes label
12  tabstat agein, statistics(N min max) by(agec)
13  tab agec, nolab
14
15
16  **#***Number of certain cancers****
17  cap use "...incidence_db0.dta", replace
18  keep patid liver pancreas endometrial
19  tempfile somec
20  save `somec', replace
21  cap use "...incidence_db1.dta", replace
22  merge 1:1 patid using `somec', update
23  keep patid gender bmirel all liver pancreas endometrial
24  groups bmirel-endometrial, missing
25  foreach var of varlist bmirel-endometrial {
26      tab `var', m
27  }
28  keep if all == 1
29  foreach var of varlist bmirel-endometrial {
30      tab `var' gender, m
31  }
32
33
34  **#***Sensitivity analyses excluding BMI<18.5 & follow-up first 2 years -- All and 4 most-common, by sex****
35  cap log close
36  cap use "  ...incidence_db1.dta", replace
37  rename agein ageatt
38
39  /*
40  stset t_all, f(all == 1) id(patid) origin(time 2)
41  strate, per(1000)
42  */
43
44  drop if t_all<2
45  replace t_all = t_all-2
46
47  mdesc
48  tab bmig, m
49  drop if bmi == .
50  drop if bmig == 0       /*removed very few subjects with BMI<18.5*/
51
52  preserve
53  keep if gender == 1
54  tempfile db1_Men
55  save `db1_Men', replace
56  restore
57
58  preserve
59  keep if gender == 2
60  tempfile db1_Women
61  save `db1_Women', replace
62  restore
63
64  foreach nm in Men Women {
65
66      foreach type in lung breast prostate colorectal all {
67
```

```
    di "-------------------------------------------------"
    di "SEX -- `nm' || TYPE -- `type' -- $S_TIME  $S_DATE"
    di "-------------------------------------------------"

    if ("`nm'" == "Men" & "`type'" == "breast") | ("`nm'" == "Women" & "`type'" == "prostate") {
        di "Few events"
    }

    else {

        qui use `db1_`nm'', clear

        stset t_`type', f(`type' == 1) id(patid)
        qui stmt, time1(df(4)) time2(df(4) start(ageatt)) timeint(t1:t2 4:4)

        preserve
        forvalues j = 0(0.1)20.1 {                      /*time since diagnosis/duration, i.e. timeinc*/
            qui clear
            qui set obs 1201
            qui generate time1 = `j'
            qui range time2 0 120 1201                  /*attained age*/
            qui tempfile temppred`n'
            qui save `temppred`n''
            local datalist `datalist' `temppred`n''
            local n = `n' + 1
        }
        qui clear
        qui set obs 0
        qui append using `datalist'
        qui tempfile timedata
        qui save `timedata'
        restore

        qui merge 1:1 _n using `timedata', nogen

        preserve
        qui predict rate, hazard time1var(time1) time2var(time2) ci
        qui gen agedd_p = time2 - time1
        qui renames time1 time2 \ diabdur_p ageatt_p
        qui order agedd_p, before(diabdur_p)
        qui keep agedd_p-ageatt_p rate*
        qui cap drop if time1 == .
        qui gen gender = "`nm'"
        qui gen type = "`type'"
        qui tempfile rates_`nm'_`type'
        qui save `rates_`nm'_`type''
        restore

        qui drop time1 time2
        qui egen float time2 = seq() in 1/6, from(5) to(10) block(1)
        qui replace time2 = time2*10
        qui gen time1_5 = 5 if time2 !=.
        qui gen time1_20 = 20 if time2 !=.
        qui predictnl lnhr  = ln(predict(hazard time1var(time1_20) time2var(time2))) - ln(predict(hazard time1var(time1_5) time2var(time2))), ci(lnhr_lci lnhr_uci)
        qui keep time2 time1_5 time1_20 lnhr*
        qui cap drop if time2 == .
        qui gen gender = "`nm'"
        qui gen type = "`type'"
        qui tempfile ratesratio_`nm'_`type'
        qui save `ratesratio_`nm'_`type''
        }
    }
}

clear
foreach nm in Men Women {
    foreach type in lung breast prostate colorectal all {
```

```
    cap append using `rates_`nm'_`type''
    }
}
foreach var of varlist rate* {
    replace `var' = `var'*1000
}
foreach var of varlist agedd_p diabdur_p ageatt_p {
    tostring `var', replace force format(%5.1f)
}
foreach var of varlist agedd_p diabdur_p ageatt_p {
    destring `var', replace
}
cap drop if rate == .
duplicates drop
cap save "...R1_Predictedrates_cancers", replace

clear
foreach nm in Men Women {
    foreach type in lung breast prostate colorectal all {
    cap append using `ratesratio_`nm'_`type''
    }
}
cap save "...R1_Predictedratesratio_cancers.dta", replace
log close


**#***Sensitivity analyses excluding BMI<18.5 & follow-up first 2 years -- BMI-related, by sex and BMI****
cap log close
cap log using "...R1_BMIcancers_log", text replace
cap use "...incidence_db1.dta", replace
rename agein ageatt
drop if bmi == .
drop if bmig == 0       /*removed very few subjects with BMI<18.5*/
tab bmig, gen(dbmi)
tabstat bmi if dbmi1 == 1, statistics(min max)

drop if t_all<2
replace t_all = t_all-2

preserve
keep if gender == 1
tempfile db1_Men
save `db1_Men', replace
restore

preserve
keep if gender == 2
tempfile db1_Women
save `db1_Women', replace
restore

foreach nm in Men Women {

        di "SEX -- `nm' -- $S_TIME  $S_DATE"

        qui use `db1_`nm'', clear

        stset t_bmirel, f(bmirel == 1) id(patid)
        stmt dbmi2 dbmi3 dbmi4 dbmi5,                                       ///
        time1(df(4) tvc(dbmi2 dbmi3 dbmi4 dbmi5) dftvc(4))                  ///
        time2(df(4) tvc(dbmi2 dbmi3 dbmi4 dbmi5) dftvc(4) start(ageatt))    ///
        timeint(t1:t2 4:4)

        preserve
        forvalues j = 0(0.1)20.1 {                      /*time since diagnosis/duration, i.e. timeinc*/
            qui clear
            qui set obs 1201
```

```
            qui generate time1 = `j'
            qui range time2 0 120 1201                  /*attained age*/
            qui tempfile temppred`n'
            qui save `temppred`n''
            local datalist `datalist' `temppred`n''
            local n = `n' + 1
        }
        qui clear
        qui set obs 0
        qui append using `datalist'
        qui tempfile timedata
        qui save `timedata'
        restore

        qui merge 1:1 _n using `timedata', nogen

        preserve
        qui predict rbmi1, hazard time1var(time1) time2var(time2) at(dbmi2 0 dbmi3 0 dbmi4 0 dbmi5 0) ci
        qui predict rbmi2, hazard time1var(time1) time2var(time2) at(dbmi2 1 dbmi3 0 dbmi4 0 dbmi5 0) ci
        qui predict rbmi3, hazard time1var(time1) time2var(time2) at(dbmi2 0 dbmi3 1 dbmi4 0 dbmi5 0) ci
        qui predict rbmi4, hazard time1var(time1) time2var(time2) at(dbmi2 0 dbmi3 0 dbmi4 1 dbmi5 0) ci
        qui predict rbmi5, hazard time1var(time1) time2var(time2) at(dbmi2 0 dbmi3 0 dbmi4 0 dbmi5 1) ci
        qui gen agedd_p = time2 - time1
        qui renames time1 time2 \ diabdur_p ageatt_p
        qui order agedd_p, before(diabdur_p)
        qui keep agedd_p-ageatt_p rbmi*
        qui cap drop if time1 == .
        qui gen gender = "`nm'"
        qui gen type = "bmirel"
        qui tempfile rates_`nm'_bmirel
        qui save `rates_`nm'_bmirel'
        restore

        qui drop time1 time2
        qui egen float time2 = seq() in 1/6, from(5) to(10) block(1)
        qui replace time2 = time2*10
        qui gen time1_5 = 5 if time2 !=.
        qui gen time1_20 = 20 if time2 !=.
        qui predictnl lnhr_bmi1  = ln(predict(hazard time1var(time1_20) time2var(time2) at(dbmi2 0 dbmi3 0 dbmi4 0 dbmi5 0))) /*
        */                       - ln(predict(hazard time1var(time1_5)  time2var(time2) at(dbmi2 0 dbmi3 0 dbmi4 0 dbmi5 0))), ci(lnhr_bmi1_lci lnhr_bmi1_uci)
        qui predictnl lnhr_bmi2  = ln(predict(hazard time1var(time1_20) time2var(time2) at(dbmi2 1 dbmi3 0 dbmi4 0 dbmi5 0))) /*
        */                       - ln(predict(hazard time1var(time1_5)  time2var(time2) at(dbmi2 1 dbmi3 0 dbmi4 0 dbmi5 0))), ci(lnhr_bmi2_lci lnhr_bmi2_uci)
        qui predictnl lnhr_bmi3  = ln(predict(hazard time1var(time1_20) time2var(time2) at(dbmi2 0 dbmi3 1 dbmi4 0 dbmi5 0))) /*
        */                       - ln(predict(hazard time1var(time1_5)  time2var(time2) at(dbmi2 0 dbmi3 1 dbmi4 0 dbmi5 0))), ci(lnhr_bmi3_lci lnhr_bmi3_uci)
        qui predictnl lnhr_bmi4  = ln(predict(hazard time1var(time1_20) time2var(time2) at(dbmi2 0 dbmi3 0 dbmi4 1 dbmi5 0))) /*
        */                       - ln(predict(hazard time1var(time1_5)  time2var(time2) at(dbmi2 0 dbmi3 0 dbmi4 1 dbmi5 0))), ci(lnhr_bmi4_lci lnhr_bmi4_uci)
        qui predictnl lnhr_bmi5  = ln(predict(hazard time1var(time1_20) time2var(time2) at(dbmi2 0 dbmi3 0 dbmi4 0 dbmi5 1))) /*
        */                       - ln(predict(hazard time1var(time1_5)  time2var(time2) at(dbmi2 0 dbmi3 0 dbmi4 0 dbmi5 1))), ci(lnhr_bmi5_lci lnhr_bmi5_uci)
        qui keep time2 time1_5 time1_20 lnhr*
        qui cap drop if time2 == .
        qui gen gender = "`nm'"
        qui gen type = "bmirel"
        qui tempfile ratesratio_`nm'_bmirel
        qui save `ratesratio_`nm'_bmirel'
}

clear
foreach nm in Men Women {
    append using `rates_`nm'_bmirel'
}
foreach var of varlist rbmi* {
    replace `var' = `var'*1000
}
foreach var of varlist agedd_p diabdur_p ageatt_p {
    tostring `var', replace force format(%5.1f)
}
foreach var of varlist agedd_p diabdur_p ageatt_p {
```

```
    destring `var', replace
}
cap drop if rbmi1 == .
duplicates drop
save "...R1_PredictedratesBMI_cancers.dta", replace

clear
foreach nm in Men Women {
    append using `ratesratio_`nm'_bmirel'
}
cap save "...R1_PredictedratesratioBMI_cancers.dta", replace
cap log close


************************************************************
**#*****[R1 FIG 1 & 2] Line plots | All cancers, by sex********
************************************************************
set scheme white_tableau
use "...Predictedrates_cancers.dta", replace
tab gender type, m

preserve
keep if gender == "Men" & type == "all"

/*[1/2]*/
twoway  (line rate ageatt_p              if agedd_p == 50, sort lcolor("247 104 161"))  ///
        (line rate ageatt_p              if agedd_p == 60, sort lcolor("221 52  151"))  ///
        (line rate ageatt_p              if agedd_p == 70, sort lcolor("174 1   126"))  ///
        (line rate ageatt_p              if agedd_p == 80, sort lcolor("122 1   119"))  ///
        (line rate_lci rate_uci ageatt_p if agedd_p == 50, sort lcolor("247 104 161" ..) lpattern(vshortdash ..) lwidth(vthin ..))  ///
        (line rate_lci rate_uci ageatt_p if agedd_p == 60, sort lcolor("221 52  151" ..) lpattern(vshortdash ..) lwidth(vthin ..))  ///
        (line rate_lci rate_uci ageatt_p if agedd_p == 70, sort lcolor("174 1   126" ..) lpattern(vshortdash ..) lwidth(vthin ..))  ///
        (line rate_lci rate_uci ageatt_p if agedd_p == 80, sort lcolor("122 1   119" ..) lpattern(vshortdash ..) lwidth(vthin ..))  ///
        , yscale(log) ylabel(2 4 8 16 32 64 128) xlabel(50(5)100, gmax)                                                             ///
        ytitle("Cancer incidence rate, per 1,000 person-years") xtitle("Attained age, years")                                       ///
        legend(rows(1) position(6) ring(0) order(1 "50 y" 2 "60 y" 3 "70 y" 4 "80 y")                                               ///
        nobox region(fcolor(none)) keygap(0.5) symxsize(4)) scale(0.8)                                                              ///
        title("[D] Rates by attained age and age at diagnosis", size(small))                                                        ///
        name("FigD", replace) nodraw

/*[1/3]*/
twoway  (line rate ageatt_p              if diabdur_p == 5,  sort lcolor("247 104 161"))    ///
        (line rate ageatt_p              if diabdur_p == 10, sort lcolor("221 52  151"))    ///
        (line rate ageatt_p              if diabdur_p == 15, sort lcolor("174 1   126"))    ///
        (line rate ageatt_p              if diabdur_p == 20, sort lcolor("122 1   119"))    ///
        (line rate_lci rate_uci ageatt_p if diabdur_p == 5,  sort lcolor("247 104 161" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
        (line rate_lci rate_uci ageatt_p if diabdur_p == 10, sort lcolor("221 52  151" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
        (line rate_lci rate_uci ageatt_p if diabdur_p == 15, sort lcolor("174 1   126" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
        (line rate_lci rate_uci ageatt_p if diabdur_p == 20, sort lcolor("122 1   119" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
        if ageatt_p>=50 & ageatt_p<=100                                                                                              ///
        , yscale(log) ylabel(1 2 4 8 16 32 64 128, gmin) xlabel(50(5)100, gmax)                                                      ///
        ytitle("Cancer incidence rate, per 1,000 person-years") xtitle("Attained age, years")                                        ///
        legend(rows(1) position(6) ring(0) order(1 "5 y" 2 "10 y" 3 "15 y" 4 "20 y")                                                 ///
        nobox region(fcolor(none)) keygap(0.5) symxsize(4)) scale(0.8)                                                               ///
        title("[B] Rates by attained age and diabetes duration", size(small))                                                       ///
        name("FigB", replace) nodraw

/*[3/1]*/
twoway  (line rate diabdur_p              if ageatt_p == 50,  sort lcolor("247 104 161"))   ///
        (line rate diabdur_p              if ageatt_p == 60,  sort lcolor("221 52  151"))   ///
        (line rate diabdur_p              if ageatt_p == 70,  sort lcolor("174 1   126"))   ///
        (line rate diabdur_p              if ageatt_p == 80,  sort lcolor("122 1   119"))   ///
        (line rate diabdur_p              if ageatt_p == 90,  sort lcolor("73  0   106"))   ///
        (line rate_lci rate_uci diabdur_p if ageatt_p == 50,  sort lcolor("247 104 161" ..)  lpattern(vshortdash ..) lwidth(vthin ..))      ///
        (line rate_lci rate_uci diabdur_p if ageatt_p == 60,  sort lcolor("221 52  151" ..)  lpattern(vshortdash ..) lwidth(vthin ..))      ///
        (line rate_lci rate_uci diabdur_p if ageatt_p == 70,  sort lcolor("174 1   126" ..)  lpattern(vshortdash ..) lwidth(vthin ..))      ///
        (line rate_lci rate_uci diabdur_p if ageatt_p == 80,  sort lcolor("122 1   119" ..)  lpattern(vshortdash ..) lwidth(vthin ..))      ///
```

```
        (line rate_lci rate_uci diabdur_p if ageatt_p == 90,  sort lcolor("73  0   106" ..)  lpattern(vshortdash ..) lwidth(vthin ..))      ///
        , yscale(log) ylabel(1 2 4 8 16 32 64 128, gmin) xlabel(0(1)20, gmax)                                                            ///
        ytitle("Cancer incidence rate, per 1,000 person-years") xtitle("Duration of diabetes, years")                                    ///
        legend(rows(1) position(6) ring(0) order(1 "50 y" 2 "60 y" 3 "70 y" 4 "80 y" 5 "90 y")                                           ///
        nobox region(fcolor(none)) keygap(0.5) symxsize(4)) scale(0.8)                                                                   ///
        title("[C] Rates by diabetes duration and attained age", size(small))                                                            ///
        name("FigC", replace) nodraw

/*[3/2]*/
twoway  (line rate diabdur_p              if agedd_p == 50, sort lcolor("247 104 161")) ///
        (line rate diabdur_p              if agedd_p == 60, sort lcolor("221 52  151")) ///
        (line rate diabdur_p              if agedd_p == 70, sort lcolor("174 1   126")) ///
        (line rate diabdur_p              if agedd_p == 80, sort lcolor("122 1   119")) ///
        (line rate_lci rate_uci diabdur_p if agedd_p == 50, sort lcolor("247 104 161" ..) lpattern(vshortdash ..) lwidth(vthin ..))     ///
        (line rate_lci rate_uci diabdur_p if agedd_p == 60, sort lcolor("221 52  151" ..) lpattern(vshortdash ..) lwidth(vthin ..))     ///
        (line rate_lci rate_uci diabdur_p if agedd_p == 70, sort lcolor("174 1   126" ..) lpattern(vshortdash ..) lwidth(vthin ..))     ///
        (line rate_lci rate_uci diabdur_p if agedd_p == 80, sort lcolor("122 1   119" ..) lpattern(vshortdash ..) lwidth(vthin ..))     ///
        , yscale(log) ylabel(2 4 8 16 32 64 128) xlabel(0(1)20)                                                                          ///
        ytitle("Cancer incidence rate, per 1,000 person-years") xtitle("Duration of diabetes, years")                                    ///
        legend(rows(1) position(6) ring(0) order(1 "50 y" 2 "60 y" 3 "70 y" 4 "80 y")                                                    ///
        nobox region(fcolor(none)) keygap(0.5) symxsize(4)) scale(0.8)                                                                   ///
        title("[A] Rates by diabetes duration and age at diagnosis", size(small))                                                        ///
        name("FigA", replace) nodraw

graph combine FigA FigB FigC FigD, cols(2) scale(0.9) xsize(6) ysize(4) name("Rates", replace) nocopies nodraw
graph save "Rates" "...R1_Fig1.gph", replace    /*R1_Fig1*/
restore

*****

preserve
keep if gender == "Women" & type == "all"

/*[1/2]*/
twoway  (line rate ageatt_p             if agedd_p == 50, sort lcolor("247 104 161"))      ///
        (line rate ageatt_p             if agedd_p == 60, sort lcolor("221 52  151"))      ///
        (line rate ageatt_p             if agedd_p == 70, sort lcolor("174 1   126"))      ///
        (line rate ageatt_p             if agedd_p == 80, sort lcolor("122 1   119"))      ///
        (line rate_lci rate_uci ageatt_p if agedd_p == 50, sort lcolor("247 104 161" ..) lpattern(vshortdash ..) lwidth(vthin ..))         ///
        (line rate_lci rate_uci ageatt_p if agedd_p == 60, sort lcolor("221 52  151" ..) lpattern(vshortdash ..) lwidth(vthin ..))         ///
        (line rate_lci rate_uci ageatt_p if agedd_p == 70, sort lcolor("174 1   126" ..) lpattern(vshortdash ..) lwidth(vthin ..))         ///
        (line rate_lci rate_uci ageatt_p if agedd_p == 80, sort lcolor("122 1   119" ..) lpattern(vshortdash ..) lwidth(vthin ..))         ///
        , yscale(log) ylabel(4 8 16 32 64) xlabel(50(5)100, gmax)                                                                        ///
        ytitle("Cancer incidence rate, per 1,000 person-years") xtitle("Attained age, years")                                            ///
        legend(rows(1) position(6) ring(0) order(1 "50 y" 2 "60 y" 3 "70 y" 4 "80 y")                                                    ///
        nobox region(fcolor(none)) keygap(0.5) symxsize(4)) scale(0.8)                                                                   ///
        title("[D] Rates by attained age and age at diagnosis", size(small))                                                             ///
        name("FigD", replace) nodraw

/*[1/3]*/
twoway  (line rate ageatt_p             if diabdur_p == 5,  sort lcolor("247 104 161"))    ///
        (line rate ageatt_p             if diabdur_p == 10, sort lcolor("221 52  151"))    ///
        (line rate ageatt_p             if diabdur_p == 15, sort lcolor("174 1   126"))    ///
        (line rate ageatt_p             if diabdur_p == 20, sort lcolor("122 1   119"))    ///
        (line rate_lci rate_uci ageatt_p if diabdur_p == 5,  sort lcolor("247 104 161" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
        (line rate_lci rate_uci ageatt_p if diabdur_p == 10, sort lcolor("221 52  151" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
        (line rate_lci rate_uci ageatt_p if diabdur_p == 15, sort lcolor("174 1   126" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
        (line rate_lci rate_uci ageatt_p if diabdur_p == 20, sort lcolor("122 1   119" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
        if ageatt_p>=50 & ageatt_p<=100                                                                                                  ///
        , yscale(log) ylabel(2 4 8 16 32 64, gmin) xlabel(50(5)100, gmax)                                                                ///
        ytitle("Cancer incidence rate, per 1,000 person-years") xtitle("Attained age, years")                                            ///
        legend(rows(1) position(6) ring(0) order(1 "5 y" 2 "10 y" 3 "15 y" 4 "20 y")                                                     ///
        nobox region(fcolor(none)) keygap(0.5) symxsize(4)) scale(0.8)                                                                   ///
        title("[B] Rates by attained age and diabetes duration", size(small))                                                            ///
        name("FigB", replace) nodraw

/*[3/1]*/
```

```
403  twoway  (line rate diabdur_p             if ageatt_p == 50,  sort lcolor("247 104 161"))   ///
404          (line rate diabdur_p             if ageatt_p == 60,  sort lcolor("221 52  151"))   ///
405          (line rate diabdur_p             if ageatt_p == 70,  sort lcolor("174 1   126"))   ///
406          (line rate diabdur_p             if ageatt_p == 80,  sort lcolor("122 1   119"))   ///
407          (line rate diabdur_p             if ageatt_p == 90,  sort lcolor("73  0   106"))   ///
408          (line rate_lci rate_uci diabdur_p if ageatt_p == 50,  sort lcolor("247 104 161" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
409          (line rate_lci rate_uci diabdur_p if ageatt_p == 60,  sort lcolor("221 52  151" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
410          (line rate_lci rate_uci diabdur_p if ageatt_p == 70,  sort lcolor("174 1   126" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
411          (line rate_lci rate_uci diabdur_p if ageatt_p == 80,  sort lcolor("122 1   119" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
412          (line rate_lci rate_uci diabdur_p if ageatt_p == 90,  sort lcolor("73  0   106" ..) lpattern(vshortdash ..) lwidth(vthin ..))       ///
413          , yscale(log) ylabel(2 4 8 16 32 64, gmin) xlabel(0(1)20, gmax)                                                                    ///
414          ytitle("Cancer incidence rate, per 1,000 person-years") xtitle("Duration of diabetes, years")                                      ///
415          legend(rows(1) position(6) ring(0) order(1 "50 y" 2 "60 y" 3 "70 y" 4 "80 y" 5 "90 y")                                             ///
416          nobox region(fcolor(none)) keygap(0.5) symxsize(4)) scale(0.8)                                                                     ///
417          title("[C] Rates by diabetes duration and attained age", size(small))                                                              ///
418          name("FigC", replace) nodraw
419
420  /*[3/2]*/
421  twoway  (line rate diabdur_p             if agedd_p == 50, sort lcolor("247 104 161"))     ///
422          (line rate diabdur_p             if agedd_p == 60, sort lcolor("221 52  151"))     ///
423          (line rate diabdur_p             if agedd_p == 70, sort lcolor("174 1   126"))     ///
424          (line rate diabdur_p             if agedd_p == 80, sort lcolor("122 1   119"))     ///
425          (line rate_lci rate_uci diabdur_p if agedd_p == 50, sort lcolor("247 104 161" ..) lpattern(vshortdash ..) lwidth(vthin ..))     ///
426          (line rate_lci rate_uci diabdur_p if agedd_p == 60, sort lcolor("221 52  151" ..) lpattern(vshortdash ..) lwidth(vthin ..))     ///
427          (line rate_lci rate_uci diabdur_p if agedd_p == 70, sort lcolor("174 1   126" ..) lpattern(vshortdash ..) lwidth(vthin ..))     ///
428          (line rate_lci rate_uci diabdur_p if agedd_p == 80, sort lcolor("122 1   119" ..) lpattern(vshortdash ..) lwidth(vthin ..))     ///
429          , yscale(log) ylabel(4 8 16 32 64) xlabel(0(1)20)                                                                                  ///
430          ytitle("Cancer incidence rate, per 1,000 person-years") xtitle("Duration of diabetes, years")                                      ///
431          legend(rows(1) position(6) ring(0) order(1 "50 y" 2 "60 y" 3 "70 y" 4 "80 y")                                                      ///
432          nobox region(fcolor(none)) keygap(0.5) symxsize(4)) scale(0.8)                                                                     ///
433          title("[A] Rates by diabetes duration and age at diagnosis", size(small))                                                          ///
434          name("FigA", replace) nodraw
435
436  graph combine FigA FigB FigC FigD, cols(2) scale(0.9) xsize(6) ysize(4) name("Rates", replace) nocopies nodraw
437  graph save "Rates" "...R1_Fig2.gph", replace
438  restore
439
440
441  ***********************************
442  **#*Sensitivity analyses***********
443  ***********************************
444
445  *Fig S7 | Absolute rates at 5 & 20 years - all comparing main vs sensitivity*
446  set scheme white_tableau
447  use "...R1_Predictedrates_cancers", clear
448  replace diabdur_p = diabdur_p + 2
449  gen analysis = "sens"
450  append using "...Predictedrates_cancers.dta"
451  replace analysis = "main" if analysis == ""
452  drop if diabdur_p>20
453  keep if inlist(ageatt_p, 50, 60, 70, 80, 90, 100)
454  keep if inlist(diabdur_p, 5, 20)
455  drop agedd_p
456  gsort gender type diabdur_p ageatt_p analysis
457  label variable ageatt_p "Age (years)"
458  label variable diabdur_p "Diabetes duration (years)"
459  order diabdur_p ageatt_p
460
461  preserve
462  replace diabdur_p = . if ageatt_p != 50
463  replace diabdur_p = . if analysis != "main"
464  replace ageatt_p  = . if analysis != "main"
465
466  forestplot rate rate_lci rate_uci if gender == "Men" & type == "all", effect("Rate") lcols(diabdur_p ageatt_p)              ///
467     nonull nonames noov nosu nowt dp(1) classic boxscale(60) astext(40) textsize(105) xlabel(0(10)180, labsize(6pt) nogrid)  ///
468     spacing(2) yline(2.5(2)22.5, lwidth(vthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(9pt))      ///
469     leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                                      ///
```

```
    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))                                  ///
    title("All cancers, Men", size(small)) name("sens_men", replace) xsize(6) ysize(4) scale(0.8) nodraw

forestplot rate rate_lci rate_uci if gender == "Women" & type == "all", effect("Rate") lcols(diabdur_p ageatt_p)            ///
    nonull nonames noov nosu nowt dp(1) classic boxscale(60) astext(40) textsize(105) xlabel(0(10)90, labsize(6pt) nogrid)   ///
    spacing(2) yline(2.5(2)22.5, lwidth(vthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(9pt))      ///
    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                                      ///
    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))                                 ///
    title("All cancers, Women", size(small)) name("sens_women", replace) xsize(6) ysize(4) scale(0.8) nodraw

graph combine sens_men sens_women, ycommon cols(2) xsize(8) ysize(6) nocopies scale(1.2) name("FigS7", replace)
graph save "FigS7" "...R1_FigS7.gph", replace
restore

*Fig S8 | Absolute rates at 5 & 20 years - most common comparing main vs sensitivity*
drop if type == "all"
replace type = proper(type)
label variable type "Cancer"
gen     ord = 1 if type == "Colorectal"
replace ord = 2 if type == "Lung"
replace ord = 3 if type == "Prostate"
replace ord = 4 if type == "Breast"
gsort gender ord diabdur_p ageatt_p analysis

order type diabdur_p ageatt_p
replace type = "" if ageatt_p != 50
replace type = "" if analysis != "main"
replace type = "" if diabdur_p == 20
replace diabdur_p = . if ageatt_p != 50
replace diabdur_p = . if analysis != "main"
replace ageatt_p  = . if analysis != "main"

forestplot rate rate_lci rate_uci if gender == "Men", effect("Rate") lcols(type diabdur_p ageatt_p)                        ///
    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(115) xlabel(0(4)40, labsize(5pt) nogrid)    ///
    spacing(2.8) yline(2.5(2)70.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))   ///
    yline(24.5 48.5, lwidth(thin) lpattern(solid)) xtitle("Rate (per 1,000 person-years)", size(7pt))                       ///
    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                                      ///
    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))                                 ///
    title("Four commonest cancers, Men", size(9pt)) name("C4_men", replace) xsize(6.5) ysize(8) scale(1.1) range(0 40) nodraw

forestplot rate rate_lci rate_uci if gender == "Women", effect("Rate") rcols(type diabdur_p ageatt_p)                      ///
    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(115) xlabel(0(2)14, labsize(5pt) nogrid)    ///
    spacing(2.8) yline(2.5(2)70.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))   ///
    yline(24.5 48.5, lwidth(thin) lpattern(solid)) xtitle("Rate (per 1,000 person-years)", size(7pt))                       ///
    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                                      ///
    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))                                 ///
    title("Four commonest cancers, Women", size(9pt)) name("C4_women", replace) xsize(6.5) ysize(8) scale(1.1) range(0 14) nodraw

graph combine C4_men C4_women, ycommon cols(2) xsize(8) ysize(6) nocopies scale(1.05) name("FigS8", replace)
graph save "FigS8" "...R1_FigS8.gph", replace

*Fig S9 | Absolute rates at 5 & 20 years - BMI-related comparing main vs sensitivity*
use "...R1_PredictedratesBMI_cancers.dta", replace
replace diabdur_p = diabdur_p + 2
gen analysis = "sens"
append using "...PredictedratesBMI_cancers.dta"
replace analysis = "main" if analysis == ""
drop if diabdur_p>20
keep if inlist(ageatt_p, 50, 60, 70, 80, 90, 100)
keep if inlist(diabdur_p, 5, 20)
drop agedd_p
gsort gender type diabdur_p ageatt_p analysis
label variable ageatt_p "Age (years)"
label variable diabdur_p "Diabetes duration (years)"

drop type
gsort gender diabdur_p ageatt_p analysis
```

```
537
538 forvalues j = 1/5 {
539     renames rbmi`j' rbmi`j'_lci rbmi`j'_uci \ rbmi_est`j' rbmi_lci`j' rbmi_uci`j'
540 }
541
542 reshape long rbmi_est rbmi_lci rbmi_uci, i(gender diabdur_p ageatt_p analysis) j(bmig)
543 order diabdur_p ageatt_p rbmi_est rbmi_lci rbmi_uci bmig gender analysis
544 gen BMI = ""
545 replace BMI = "≥18.5-<25" if bmig == 1
546 replace BMI = "≥25-<30"   if bmig == 2
547 replace BMI = "≥30-<35"   if bmig == 3
548 replace BMI = "≥35-<40"   if bmig == 4
549 replace BMI = "≥40"       if bmig == 5
550 gsort gender bmig diabdur_p ageatt_p analysis
551
552 replace diabdur_p = . if ageatt_p != 50
553 replace diabdur_p = . if analysis != "main"
554 replace ageatt_p  = . if analysis != "main"
555
556 *Men*
557 forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Men" & bmig == 1, effect("Rate") lcols(diabdur_p ageatt_p)              ///
558    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)     ///
559    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))   ///
560    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                                      ///
561    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                                      ///
562    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))                                 ///
563    title("BMI ≥18.5-<25 kg/m{superscript:2}, Men", size(9pt)) name("bmig1_men", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw
564
565 forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Men" & bmig == 2, effect("Rate") lcols(diabdur_p ageatt_p)              ///
566    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)     ///
567    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))   ///
568    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                                      ///
569    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                                      ///
570    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))                                 ///
571    title("BMI ≥25-<30 kg/m{superscript:2}, Men", size(9pt)) name("bmig2_men", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw
572
573 forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Men" & bmig == 3, effect("Rate") lcols(diabdur_p ageatt_p)              ///
574    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)     ///
575    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))   ///
576    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                                      ///
577    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                                      ///
578    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))                                 ///
579    title("BMI ≥30-<35 kg/m{superscript:2}, Men", size(9pt)) name("bmig3_men", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw
580
581 forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Men" & bmig == 4, effect("Rate") lcols(diabdur_p ageatt_p)              ///
582    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)     ///
583    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))   ///
584    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                                      ///
585    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                                      ///
586    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))                                 ///
587    title("BMI ≥35-<40 kg/m{superscript:2}, Men", size(9pt)) name("bmig4_men", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw
588
589 forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Men" & bmig == 5, effect("Rate") lcols(diabdur_p ageatt_p)              ///
590    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)     ///
591    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))   ///
592    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                                      ///
593    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                                      ///
594    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))                                 ///
595    title("BMI ≥40 kg/m{superscript:2}, Men", size(9pt)) name("bmig5_men", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw
596
597 graph combine bmig1_men bmig2_men bmig3_men bmig4_men bmig5_men, ycommon cols(5) xsize(8) ysize(6) nocopies scale(1.05) name("FigS9_men", replace)
598
599 *Women*
600 forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Women" & bmig == 1, effect("Rate") lcols(diabdur_p ageatt_p)            ///
601    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)     ///
602    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))   ///
603    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                                      ///
```

```
    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                  ///
    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))             ///
    title("BMI ≥18.5-<25 kg/m{superscript:2}, Women", size(9pt)) name("bmig1_women", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw

forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Women" & bmig == 2, effect("Rate") lcols(diabdur_p ageatt_p)           ///
    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)    ///
    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))  ///
    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                  ///
    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                  ///
    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))             ///
    title("BMI ≥25-<30 kg/m{superscript:2}, Women", size(9pt)) name("bmig2_women", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw

forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Women" & bmig == 3, effect("Rate") lcols(diabdur_p ageatt_p)           ///
    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)    ///
    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))  ///
    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                  ///
    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                  ///
    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))             ///
    title("BMI ≥30-<35 kg/m{superscript:2}, Women", size(9pt)) name("bmig3_women", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw

forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Women" & bmig == 4, effect("Rate") lcols(diabdur_p ageatt_p)           ///
    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)    ///
    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))  ///
    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                  ///
    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                  ///
    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))             ///
    title("BMI ≥35-<40 kg/m{superscript:2}, Women", size(9pt)) name("bmig4_women", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw

forestplot rbmi_est rbmi_lci rbmi_uci if gender == "Women" & bmig == 5, effect("Rate") lcols(diabdur_p ageatt_p)           ///
    nonull nonames noov nosu nowt dp(1) classic boxscale(35) astext(45) textsize(85) xlabel(0(4)40, labsize(7pt) nogrid)    ///
    spacing(2.2) yline(2.5(2)22.5, lwidth(vvthin) lpattern(vshortdash)) xtitle("Rate (per 1,000 person-years)", size(7pt))  ///
    xtitle("Rate (per 1,000 person-years)", size(7pt))                                                  ///
    leftjustify ciopts(lwidth(vthin)) plotid(analysis)                                                  ///
    box1opts(mcolor(red)) ci1opts(lcolor(red)) box2opts(mcolor(blue)) ci2opts(lcolor(blue))             ///
    title("BMI ≥40 kg/m{superscript:2}, Women", size(9pt)) name("bmig5_women", replace) xsize(6.5) ysize(8) scale(0.7) range(0 40) nodraw

graph combine bmig1_women bmig2_women bmig3_women bmig4_women bmig5_women, ycommon cols(5) xsize(8) ysize(6) nocopies scale(1.05) name("FigS9_women", replace)

graph combine FigS9_men FigS9_women, ycommon rows(2) xsize(8) ysize(6) nocopies scale(1.05) name("FigS9", replace)
graph save "FigS9" "...R1_FigS9.gph", replace
graph close _all
```