eAppendix 1. Hazard ratios for continuous duration-based measures in Health 2000 and NHANES, ages 50-74 at survey

	Health 2000 Finland			NHANES[1] United States		
Model	*HR*	*95% CI*	*Coefficient/ SE*	*HR*	*95% CI*	*Coefficient/ SE*
Obesity Duration[2], years	1.020	(1.008, 1.032)	3.381	1.016	(1.010, 1.022)	5.298
Obesity Duration[3], percentage	1.010	(1.004, 1.015)	3.654	1.007	(1.004,1.009)	5.350
Years Since Becoming Obese[4]	1.018	(1.008, 1.028)	3.606	1.015	(1.010, 1.021)	5.617
BMI-years above 30.0[5]	1.004	(1.002, 1.007)	3.583	1.002	(1.001,1.003)	5.260

Note: Obesity is defined as a BMI≥30.0. Hazard ratios are for a 1-unit change in each measure. Duration-based measures estimated from retrospective weight at earlier ages and measured weight at time of survey. Health 2000 models adjusted for attained age, sex, educational attainment, and cigarette smoking. Duration variables in Health 2000 calculated from information on BMI at ages 20, 30, 40, and 50 and at baseline. Deaths were followed through December 31, 2011 (deaths=335). Estimates reflect sample weighting.

[1] From Table 6 in Preston et al. based on data from NHANES 3 (1988-1994) and the 1999-2004 continuous waves. Duration measures calculated from retrospectively reported weight at age 25, age 10 years prior to survey, and measured weight at time of survey. Models adjusted for age at survey, race/ethnicity, educational attainment, and cigarette smoking. Deaths were followed through December 31st, 2004.

[2] Count of years spent obese between age 20 (Health 2000) or age 25 (NHANES) and age at time of survey

[3] Percentage of life spent obese between age 20 (Health 2000) or age 25 (NHANES) and age at time of survey.

[4] Count of years since first becoming obese since age 20 (Health 2000) or age 25 (NHANES).

[5] Combined measure of number of years obese multiplied by the mean number of BMI units above a BMI of 30.0.

CI: Confidence Intervals
HR: Hazard Ratios
SE: Standard Error

eAppendix 2. Cross-tabulation of categorical BMI at survey and historical peak BMI, ages 50-74 at survey

	Survey BMI			
Historical Peak BMI	BMI<25	Overweight (25.0-29.9)	Obese (≥30.0)	*Total*
BMI<25	579 (82)	314 (29)	40 (6)	933 (37)
Overweight	119 (17)	684 (63)	318 (45)	1,125 (45)
Obese	6 (1)	92 (8)	353 (50)	451 (18)
Total	704 (100)	1,090 (100)	711 (100)	2,505 (100)

Note: Cells show sample sizes (n) and column percentages (in parentheses). Survey BMI is BMI at time of survey in 2000-2001. Historical peak BMI is the maximum BMI observed from information on BMI at ages 20, 30, 40, 50. Historical peak BMI excludes information on BMI at time of survey.

Source: Health 2000 Finland

eAppendix 3. Data acquisition and sample Stata code.

The Health 2000 data are available to researchers by application to the Health 2000 research group at the National Institute of Health and Welfare (THL) in Finland. The application requires a one-page research proposal and signed agreement letter. Further information about the data are available on-line (www.terveys2000.fi).

```
                    /**** HEALTH 2000 DATA—Analytical Code ****/

                                /* Data Set up */
clear
clear matrix
set mem 1g

* Set Global Directories
global dir ""
global data ""

* Import data
use "$dir$data\H2000.dta", replace



                        /* Variable Edits & Construction */

*** BMI Variable Construction

       * Weight Hx non-missing Indicator
       g weightindic=bmi20<. & bmi30<. & bmi40<. & bmi50<. & bmi<.

       * Categorical BMI at Survey
       g bmicat_s=bmicat
       replace bmicat_s=30 if bmicat==35

       * Linear BMI at Survey
       g bmi_s=bmi
       replace bmi_s=bmi-25 if bmi>=25 & bmi<.
       replace linbl=0 if bmi<25 & bmi<.

       * Retrospective BMI--set up
       g lin20=bmi20 /* age 20 */
       replace lin20=bmi20-25 if bmi20>=25 & bmi20<.
       replace lin20=0 if bmi20<25 & bmi20<.

       g lin30=bmi30
       replace lin30=bmi30-25 if bmi30>=25 & bmi30<.
       replace lin30=0 if bmi30<25 & bmi30<.

       g lin40=bmi40
       replace lin40=bmi40-25 if bmi40>=25 & bmi40<.
       replace lin40=0 if bmi40<25 & bmi40<.

       g lin50=bmi50
       replace lin50=bmi50-25 if bmi50>=25 & bmi50<.
       replace lin50=0 if bmi50<25 & bmi50<.

       * Overall and historical peak BMI--linear
       egen bmi_o=rowmax(lin20 lin30 lin40 lin50 linbl) if lin20<. & lin30<. &
       lin40<. & lin50<. & linbl<.
```

```
egen bmi_h=rowmax(lin20 lin30 lin40 lin50) if lin20<. & lin30<. &  lin40<. &
lin50<. & linbl<.

* Overall and historical peak BMI--categorical
* Overall
g bmicat_o=.
replace bmicat_o=0 if bmi_o==0 & bmi_o<.
replace bmicat_o=1 if bmi_o~=0 & OBlinmax1==0 & bmi_o<.
replace bmicat_o=2 if OBlinmax1>0 & bmi_o<.

* Historical
g bmicat_h=.
replace bmicat_h=0 if bmi_h==0 & bmi_h<.
replace bmicat_h=1 if bmi_h~=0 & OBlinmax2==0 & bmi_h<.
replace bmicat_h=2 if OBlinmax2>0 & bmi_h<.


* CTS simple duration and pack-years
*** BMI=25 threshold
* centered BMI on 25
cap drop cbmi*
g cbmi20=bmi20-25
g cbmi30=bmi30-25
g cbmi40=bmi40-25
g cbmi50=bmi50-25
g cbmi=bmi-25

* interval
cap drop interval*
forvalues i=1/4 {
g interval`i'=.
}

/* Age 20 and Age 30 */
replace interval1=1 if cbmi20<0 & cbmi30<0
/* not obese in either period */
replace interval1=2 if cbmi20>=0 & cbmi30>=0
/* obese in both periods */
replace interval1=3 if cbmi20<0 & cbmi30>=0                /* gainers */
replace interval1=4 if cbmi20>=0 & cbmi30<0                /* losers */

/* Age 30 and Age 40 */
replace interval2=1 if cbmi30<0 & cbmi40<0           /* not obese in either
period */
replace interval2=2 if cbmi30>=0 & cbmi40>=0   /* obese in both periods */
replace interval2=3 if cbmi30<0 & cbmi40>=0                /* gainers */
replace interval2=4 if cbmi30>=0 & cbmi40<0                /* losers */

/* Age 40 and Age 50 */
replace interval3=1 if cbmi40<0 & cbmi50<0    /* not obese in either period */
replace interval3=2 if cbmi40>=0 & cbmi50>=0               /* obese in both
periods */
replace interval3=3 if cbmi40<0 & cbmi50>=0                /* gainers */
replace interval3=4 if cbmi40>=0 & cbmi50<0                /* losers */

/* Age 50 and BL */
replace interval4=1 if cbmi50<0 & cbmi<0       /* not obese in either period */
replace interval4=2 if cbmi50>=0 & cbmi>=0     /* obese in both periods */
replace interval4=3 if cbmi50<0 & cbmi>=0                  /* gainers */
replace interval4=4 if cbmi50>=0 & cbmi<0                  /* losers */
```

```
*** Simple Obesity Duration ***
cap drop _25obdur*
g _25obdur=.
g _25obdur1=.
g _25obdur2=.
g _25obdur3=.
g _25obdur4=.

/* Interval 1 */
tempvar early recent l
g `early'=cbmi20
g `recent'=cbmi30
g `l'=10
replace _25obdur1=0 if interval1==1
replace _25obdur1=`l' if interval1==2
replace _25obdur1=(`l'*`recent')/(`recent'-`early') if interval1==3
replace _25obdur1=(`l'*`early')/(`early'-`recent') if interval1==4

/* Interval 2 */
tempvar early recent l
g `early'=cbmi30
g `recent'=cbmi40
g `l'=10
replace _25obdur2=0 if interval2==1
replace _25obdur2=`l' if interval2==2
replace _25obdur2=(`l'*`recent')/(`recent'-`early') if interval2==3
replace _25obdur2=(`l'*`early')/(`early'-`recent') if interval2==4

/* Interval 3 */
tempvar early recent l
g `early'=cbmi40
g `recent'=cbmi50
g `l'=10
replace _25obdur3=0 if interval3==1
replace _25obdur3=`l' if interval3==2
replace _25obdur3=(`l'*`recent')/(`recent'-`early') if interval3==3
replace _25obdur3=(`l'*`early')/(`early'-`recent') if interval3==4

/* Interval 4 */
tempvar early recent l
g `early'=cbmi50
g `recent'=cbmi
g `l'=age-50

replace _25obdur4=0 if interval4==1
replace _25obdur4=`l' if interval4==2
replace _25obdur4=(`l'*`recent')/(`recent'-`early') if interval4==3
replace _25obdur4=(`l'*`early')/(`early'-`recent') if interval4==4

replace _25obdur=_25obdur1+_25obdur2+_25obdur3+_25obdur4
replace _25obdur=. if weightindic~=1
rename _25obdur obdur

*** Pack-Years
cap drop _25packyrs*
g _25packyrs=.
g _25packyrs30=. /* since age 30 */
g _25packyrs1=.
g _25packyrs2=.
g _25packyrs3=.
g _25packyrs4=.
```

```
        /* Interval 1 */
        tempvar early recent l
        g `early'=cbmi20
        g `recent'=cbmi30
        g `l'=10

        replace _25packyrs1=0 if interval1==1
        replace _25packyrs1=((`early'+`recent')/2)*`l' if interval1==2
        replace _25packyrs1=(_25obdur1*`recent')/2 if interval1==3
        replace _25packyrs1=(_25obdur1*`early')/2 if interval1==4

        /* Interval 2 */
        tempvar early recent l
        g `early'=cbmi30
        g `recent'=cbmi40
        g `l'=10

        replace _25packyrs2=0 if interval2==1
        replace _25packyrs2=((`early'+`recent')/2)*`l' if interval2==2
        replace _25packyrs2=(_25obdur2*`recent')/2 if interval2==3
        replace _25packyrs2=(_25obdur2*`early')/2 if interval2==4

        /* Interval 3 */
        tempvar early recent l
        g `early'=cbmi40
        g `recent'=cbmi50
        g `l'=10

        replace _25packyrs3=0 if interval3==1
        replace _25packyrs3=((`early'+`recent')/2)*`l' if interval3==2
        replace _25packyrs3=(_25obdur3*`recent')/2 if interval3==3
        replace _25packyrs3=(_25obdur3*`early')/2 if interval3==4

        /* Interval 4 */
        tempvar early recent l
        g `early'=cbmi50
        g `recent'=cbmi
        g `l'=age-50

        replace _25packyrs4=0 if interval4==1
        replace _25packyrs4=((`early'+`recent')/2)*`l' if interval4==2
        replace _25packyrs4=(_25obdur2*`recent')/2 if interval4==3
        replace _25packyrs4=(_25obdur2*`early')/2 if interval4==4

        replace _25packyrs=_25packyrs1+_25packyrs2+_25packyrs3+_25packyrs4
        replace _25packyrs=. if weightindic~=1
        rename _25packyrs packyrs

                        /* Analytical Sample Identification */

* Sample identification
        cap drop sample
        g sample=(minifin==0 & insample==1 &  KYSELY1_OLO2==1)
        /* KYSELY1_OLO2 is an indicator for answering structured home interview */
        * Subpopulation identification 1--non-missing on all covariates
        g subpop=sample & bmi20<. & bmi30<. & bmi40<. & bmi50<. & bmi<. & death13~=. &
        never<. & educ3<. & aweight~=. & aweight~=0 & BD07~=1
        * Subpopulation identification 2--non-missing on all covariates plus
        elimination of baseline underweight and outliers; age restriction
        g asamp=subpop & _30packyrs<500 & under==0 & OBlinmax2<=25 & OBlinbl<=25 &
        age>=50 & age<75
```

```
                                /* RESULTS */

** Table 1:
svyset [pw=aweight]
svy, subpop(asamp): mean obdur packyrs
svy, subpop(asamp): proportion bmicat_s /* survey bmi */
svy, subpop(asamp): proportion bmicat_o /* overall peak bmi */
svy, subpop(asamp): mean bmi_s bmi_o bmi_h /* bmi_h is historical peak bmi */

* Build Person-year format
stset datecensor13 [pw=aweight], failure(death13) origin(time examdate)
if(asamp) id(id) scale(30.5)
stsplit dur, every(12)
sort id _t0
by id: g age_tv=age+_n-1

** Table 2
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 obdur
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 packyrs

** Table 3
svy, subpop(asamp): stcox age_tv female i.educ3 i.bmicat_o
svy, subpop(asamp): stcox age_tv female i.educ3 i.bmicat_s

** Table 4
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 bmi_s
/* model 1 */
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 bmi_o
/* model 2 */
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 bmi_h
/* model 3 */
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 bmi_s bmi_h
/* model 4 */


** Appendix Table 1
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 obdur_30
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 obper_30
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 yrsince_30
svy, subpop(asamp): stcox age_tv female i.educ3 i.smoke2 packyrs_30

** Appendix Table 2
/** Estimated prior to building person-year file **/
tab bmicat_s bmicat_h if asamp, col
```