```
capture log close
clear
log using  "p:\ocs\results\an_1.ovarian.new.doc", text replace

* log using  "c:\usr\rino\projects\ovarian_new\results\an_
1.ovarian.new.doc", text replace



/*
********************************************************
Date        :  1/07/2005
Update      :  1/07/2005
Authors     :  Rino , Asa


Name of file:      an_1.ovarian.new.do
Input Data file:   ovariannew_1.dta
Output Data file:  ovariannew_2.dta
Note:              it calls sub-programs

Content:    1) This file will produce the data set that we will use
in the ongoing projects
            2) It creates exposure information
            4) It creates outcome information


*********************************************************

*/




/* Step 1 */


clear
set mem 40m
use p:\ocs\data\merged\ovariannew_1, clear
*use  "c:\usr\rino\projects\ovarian_new\data\merged\ovariannew_1",
clear


/*
*********************************************************************
****************
* GENERAL EXCLUSIONS
*********************************************************************
****************
*/

drop if exclude==1
drop if lopnr==11999

* this case was excluded due to the presence of colon cancer
* Rino: who decided so?


/* Step 2 */

/* value labels*/
```

```
label define yesnno  0 "No"  1"Yes"



/* Step 3 */

*********************************************************************
****************
*     TYPES OF TUMOURS:
*********************************************************************
****************
/* How many gave answer on malbor malmal */

tab na0704, mis

gen malig=.
replace malig=1 if malbor==1
replace malig=2 if malmal==1
label var malig "Malignancy: Invasive/Borderline"
label var malmal "Invasive tumors"
label var malbor "Borderline tumors"
label define malig 1 "borderline" 2"invasive", nofix
label values malig malig
tab malig, mis

replace hist_0=6 if hist_0==99
replace hist_0=6 if hist_0==199
replace hist_0=6 if hist_0==299
label var hist_0 "tumor histological subtype"
label define hist_0 1"serous" 2"mucinous" 3"endometrioid"
4"clearcell" 5"undifferentiated" 6"others", nofix
label values hist_0 hist_0
tab hist_0, mis

/* there are 23 missing !!! */


*********************************************************************
**********************
*     COVARIATES:
*********************************************************************
**********************


** Reason for diagnosis

gen q3=.
replace q3=0 if na0301==1
replace q3=1 if gyninv==1
replace q3=2 if symptom==1
replace q3=3 if other!=""
label var q3 "reason for diagnosis"
label define q3 0"not stated" 1"routine" 2"symptoms" 3"other", nofix
label values q3 q3
tab q3, mis

** Age at diagnosis

* generate a birth date
capture gen y=substr(pnr,1,4)
capture gen m=substr(pnr,5,2)
capture gen d=substr(pnr,7,2)
capture gen newdmy=d+"-"+m+"-"+y
```

```
gen birthdate = date(newdmy,"dmy")
format birthdate %d
format birthdate %dD_m_Y

* generate age at diagnosis from the birthdate
gen diagage=(diagdat-birthdate)/365.2

gen diagage_cat=.
replace diagage_cat=1 if diagage>=50 & diagage<55
replace diagage_cat=2 if diagage>=55 & diagage<60
replace diagage_cat=3 if diagage>=60 & diagage<65
replace diagage_cat=4 if diagage>=65 & diagage<70
replace diagage_cat=5 if diagage>=70 & diagage<75

* two women 49 years old were considered 50

replace diagage_cat=1 if diagage < 50
label var diagage_cat "Categorical age at diagnosis"
label define diagage_cat 1"50 - 54 yrs" 2"55 - 59 yrs" 3"60 - 64
yrs" 4"65 - 69 yrs" 5"70 - 75 yrs", modify
label values diagage_cat diagage_cat
tab  diagage_cat, mis



** FIGO **

* Creating a variable for Figo Staging.  Will create a new variable
called flag figo
* to see if the total sum is >1 to see if more than 1 box is checked.
destring figoix, gen(figo1x)
destring figoiix, gen(figo2x)
destring figoiiix, gen(figo3x)
destring figoivx, gen(figo4x)
replace figo2x=1 if figo2x==2
replace figo3x=1 if figo3x==3
replace figo4x=1 if figo4x==4
egen flagfigo = rsum ( na0703 figoia figoib figoic figo1x figoiia
figoiib figoiic figo2x figoiiia figoiiib figoiiic figo3x figo4x)
tab flagfigo, mis
* Each case  has only one classification of figo staging

gen figo=.
replace figo=1 if figoia==1
replace figo=2 if figoib==1
replace figo=3 if figoic==1
replace figo=4 if figoix!=""
replace figo=5 if figoiia==1
replace figo=6 if figoiib==1
replace figo=7 if figoiic==1
replace figo=8 if figo2x==1
replace figo=9 if figoiiia==1
replace figo=10 if figoiiib==1
replace figo=11 if figoiiic==1
replace figo=12 if figo3x==1
replace figo=13 if figo4x==1
label var figo "Stage of disease wrt Figo"
label define figo 1"Ia" 2"Ib" 3"Ic" 4"Ix" 5"IIa" 6"IIb" 7"IIc"
8"IIx" 9"IIIa" 10"IIIb" 11"IIIc" 12"IIIx" 13"IV", modify
label values figo figo
tab figo, mis

gen figo_cat=.
replace figo_cat=1 if figo>=1 & figo<=4
```

```
replace figo_cat=2 if figo>=5 & figo<=8
replace figo_cat=3 if figo>=9 & figo<=12
replace figo_cat=4 if figo>=13
label var figo_cat "collapsed figo stages"
label define figo_cat 1"I" 2"II" 3"III" 4"IV", modify
label values figo_cat figo_cat
tab figo_cat, mis
tab figo figo_cat, mis


** Grade of differentiation
 gen who=.
replace who=0 if na0706==1
replace who=1 if whohi==1
replace who=2 if whome==1
replace who=3 if wholo==1
label var who "grade of differentiation according to WHO"
label define who 0"not stated" 1"highly differentiated" 2"medium
differentiated" 3"lowly differentiated", modify
label values who who
tab who, mis



** Residual tumor size occasion 1

/* To be checked carefully:
This is not documented by CM and reclassification is based on
id Stata number !!!!
*/



gen residual1=.
replace residual1=2 if resnm1==1
replace residual1=3 if na0603_1==1
replace residual1=1 if rescm1>=2 & rescm1!=.
replace residual1=0 if rescm1<2 & rescm1!=.
replace residual1=0 if restxt1=="<2cm"
replace residual1=1 if restxt1==">2cm"
replace residual1=1 if restxt1=="2 kubikcm"
replace residual1=0 if restxt1=="<1cm"
replace residual1=1 if restxt1=="2-3 kubikcm"
replace residual1=0 if restxt1=="några mm"
replace residual1=1 if restxt1=="ett område 4x5cm"
replace residual1=1 if restxt1=="5mmx5cm parti"
replace residual1=0 if restxt1=="mindre än 2 cm i bäckenperitoneum"
replace residual1=0 if restxt1=="tunnt btäm av tumormassa 1-2 cm"
replace residual1=0 if restxt1=="''smärre kvarvarande tumörbitar"
replace residual1 = 0 if lopnr==11118

* edit lopnr residual1 if lopnr==11118
* replace residual1 = 0 in 101

replace residual1=0 if restxt1=="hasselnötsstor"
replace residual1=0 if restxt1=="brunbönstor"
replace residual1=0 if restxt1=="hasselnöt"
replace residual1=0 if restxt1=="risgryn"
replace residual1=0 if restxt1=="knappnålshuvud"
replace residual1=0 if restxt1=="brunböna"
replace residual1=0 if restxt1=="Brunböna"
replace residual1=0 if restxt1=="ärtstora"
replace residual1=0 if restxt1=="spelkula"
replace residual1=0 if restxt1=="pepparkorn"
```

```
replace residual1=0 if restxt1=="minimal disease"

replace residual1 = 0 if lopnr==11781
*edit lopnr residual1 if lopnr==11781
*replace residual1 = 0 in 603

replace residual1=0 if restxt1=="liten"

replace residual1 = 0 if lopnr==11799
*edit lopnr residual1 if lopnr==11799
*replace residual1 = 0 in 613

replace residual1=1 if restxt1=="tumände"
replace residual1=1 if restxt1=="tunn matta av peritonealcarcinos"
replace residual1=1 if restxt1=="tumändstor"
replace residual1=1 if restxt1=="tennisboll"
replace residual1=1 if restxt1=="stor tumörbörda"
replace residual1=1 if restxt1=="stora tumörmassor"
replace residual1=1 if restxt1=="valnötstor"

replace residual1 = 1 if lopnr==11245
*edit lopnr residual1 if lopnr==11245
*replace residual1 = 1 in 197

replace residual1=1 if restxt1=="pingpongboll"
replace residual1=1 if restxt1=="barnhandflata"
replace residual1=1 if restxt1=="stora tumörer"
replace residual1=1 if restxt1=="dubbelt bönstor"
replace residual1=1 if restxt1=="jaffaapelsin"
replace residual1=1 if restxt1=="tumör i hela buken"
replace residual1=1 if restxt1=="halv knytnäve"
replace residual1=1 if restxt1=="kokosnöt"
replace residual1=1 if restxt1=="mandarin"
replace residual1=1 if restxt1=="1/2 x 5 x 5 cm"
replace residual1=1 if restxt1=="handflata"
replace residual1=1 if restxt1=="stora tumörmassor"
replace residual1=1 if restxt1=="hönsägg"
replace residual1=0 if restxt1=="<2cm spritt"
replace residual1=0 if restxt1=="5 krona"
replace residual1=0 if restxt1=="körsbär"
replace residual1=0 if restxt1=="mindre än 1 cm"
replace residual1=0 if restxt1=="smärre kvarvarande tumörbitar"
replace residual1=0 if restxt1=="knottrot <2cm"
replace residual1=0 if restxt1=="ärtstor"
replace residual1=0 if restxt1=="några mm-stora knottror"
replace residual1=0 if restxt1=="brun böna"
replace residual1=0 if restxt1=="1/2 cm tjock matta"
replace residual1=0 if restxt1=="mmtjock belägnn.  I dubbel
tumändstort område"

replace residual1 = 0 if lopnr==11927
*edit lopnr residual1 if lopnr==11927
*replace residual1 = 0 in 421

label var residual1 "residual tumor size occasion 1"
label define residual1 0"< 2cm" 1"> 2cm" 2"non-measurable" 3"not
stated", nofix
label values residual1 residual1


* opy1
label var opy1 "Operation: primary surgery"
label define opy1 0"no" 1"yes", nofix
label values opy1 opy1
```

```
gen rad = .
replace rad=0 if radn==1
replace rad=1 if rady==1
replace rad=2 if na0901==1
label var rad "radiation therapy line 1"
label define rad 0"no" 1"yes" 2"not stated", nofix
label values rad rad

label var chemoy1 "chemotherapy line 1"
label values chemoy1 opy1

gen endo = .
replace endo=1 if eny==1
replace endo=0 if eny==0
replace endo=2 if na1001==1
replace endo=0 if enstd>=lastd1 & endo!=2
* (from Kjell's review of the case numbers and definition of
endocrine therapy to determine accurate endocrine therapy)
label var endo "endocrine therapy"
label values endo rad


/* Social Status */

gen myses=0
replace myses=1 if ses >3
replace myses=. if ses ==.


/* BMI */

/*
f4= weight
f2= height
*/
gen f2_cm=f2/100
gen bmi=f4/(f2_cm*f2_cm)
label var bmi "body mass index 1 yr ago"
codebook bmi

gen bmi_cat = bmi
replace bmi_cat = 1 if bmi<18.5
replace bmi_cat = 2 if bmi>=18.5 & bmi<25 & bmi!=.
replace bmi_cat = 3 if bmi>=25 & bmi<30 & bmi!=.
replace bmi_cat = 4 if bmi>=30 & bmi!=.
label var bmi_cat "body mass index categorical"
label define bmi_cat 1 "underweight" 2 "normal" 3 "overweight" 4
"obese", nofix
label values bmi_cat bmi_cat
codebook bmi_cat


/* Smoking */

* smoy (from Tomas' data)
label var smoy "smoking one yr ago from questionnaires"
label define smoy 1 "yes" 2"no", nofix
label values smoy smoy
codebook smoy


gen smoke=0 if sm==2
replace smoke =1 if smoy==1
```

```
replace smoke =2 if smoy==2

/* Hormonal Contraception */

* f43 (from Tomas' data)
label var f43 "hormonal contraceptives ever"
label define f43 1 "no" 2 "yes", nofix
label values f43 f43
codebook f43

* occombdu (from Tomas' data)
label var occombdu "any combined OC's duration (days)"
codebook occombdu


/* Age at menarche */
* f37 (from Tomas' data)
label var f37 "age at menarche"
codebook f37


gen menage_13=0
replace menage_13=1 if f37 >13 & f37 <.

label variable menage_13 "Menar Age >13"






/* Parity */

* para0 (from Tomas' data)
replace para0=1-para0
label var para0 "parity categorical"
label define para0 0 "NO" 1 "YES", nofix
label values para0 para0
codebook para0

* parity (from Tomas' data)
label var parity "number of full term pregnancies"
codebook parity

gen parnew=0
replace parnew=1 if parity >0
replace parnew=2 if parity >2
replace parnew=3 if parity >4
replace parnew=. if parity ==.

label variable parnew "N.of Children"
label define parfmt 0 "0"  1 "1-2" 2 "3-4" 3">=5"
label val parnew parfmt



* occombdu (from Tomas' data)
label var occombdu "any combined OC's duration (days)"
codebook occombdu

* mpage (from Tomas' data)
label var mpage "age at menopause"
codebook mpage
```

```
gen mpage_50=0
replace mpage_50=1 if mpage >50 & mpage <.
label variable mpage_50 "Menop Age >50"
label val  mpage_50  yesno


* f32_a (from Tomas' data)
label var f32_a "tubal ligation ever"
label define f32_a 0 "no" 1"yes", nofix
label values f32_a f32_a
codebook f32_a

* f35 (from Tomas' data)
label var f35 "ovarian cancer in mother/sister"
label define f35 1"yes" 2"no" 3"don't know", nofix
label values f35 f35



cd p:\ocs\\programs
*cd "C:\usr\rino\projects\ovarian_new\programs"

do cr.hrt.do
do cr.surv.new.do


save  "p:\ocs\data\merged\ovariannew_2", replace
      save  "c:\usr\rino\projects\ovarian_new\data\merged\ovariannew_
2", replace




log close
```