

```
capture log close
clear
log using "c:\usr\rino\projects\ovarian_new\results\cr.tomas.doc",
text replace
```

```
/*
*****
```

```
Date       : 24/01/2005
Update     : 27/02/2006
Authors    : Rino Bellocco
```

```
Name of file: cr.tomas.do
Input Data file : epct_04.dta, pnr.tab
Output Data file: tomas.dta, pnr_1
```

```
Subject:      This file does:
               1: Reads Tomas' main data file epct_04.dta
               2: Create pnr1.dta from pnr.tab : it has the lopnr and
pnr
               2: what CM has used is WRONG but results are ok ???
               3: merge epct_04 and pnr1.dta
               4: select cases and deletes the duplicate case with
duplicate pnr
               5: 193004061309 and lopnr 12026 (deleted)
               6: keeps the variables needed for analysis 1 and saves
this file as tomas.dta
               7: Check PNR: 5 wrong
               8: add two more variables (le_vgtr le_potr)
```

	pnr	lopnr
63.	194107041043	11069
Right:	194107049043	
92.	193505068947	11107
Right:	193205068947	
108.	194212284609	11126
Right:	194312284604	
166.	191910261665	11195
Right:	191910261666	
232.	193401073327	11279
Right:	193401073329	

```
*****
```

```
*/
set mem 40m
```

```
*****
* 1
*****
```

```
use "C:\usr\rino\projects\ovarian_new\data\tomas\epct_04.dta", clear
sort lopnr
save "C:\usr\rino\projects\ovarian_new\data\tomas\epct_04.dta",
replace
```

```

clear
insheet lopnr pnr using "C:\usr\rino\projects\ovarian_new\data\tomas
\pnr.tab", double
format pnr %10.0f
tostring pnr, gen(newpnr)
gen var1=substr(newpnr,1,6)
gen var2=substr(newpnr,7,4)
gen newpnrl="19" + var1+ var2
drop var1 var2
drop pnr newpnr
rename newpnrl pnr
sort lopnr
save "C:\usr\rino\projects\ovarian_new\data\tomas\pnrl.dta",replace

```

```

*****
*   II and III
*****

```

```

use "C:\usr\rino\projects\ovarian_new\data\tomas\pnrl.dta", clear
sort lopnr
merge lopnr using "C:\usr\rino\projects\ovarian_new\data\tomas\epct_
04.dta"
tab _merge

```

```

/* Take only the study subjects */
keep if _merge==3
drop _merge

```

```

*****
*   IV
*****

```

```

/* Take the cases only */
keep if case==1

```

```

/* Chantal : why?? */
* drop in 814
/* There are two women with different lopnr */

```

```

drop if lopnr==12026

```

```

*****
*   V
*****

```

```

keep lopnr pnr ch* agefb agelb bfdur eodu eotr ep_ctdu ep_cttr
ep_cydu ep_cytr f2 f4 sm f35 f43 f49 f10tel ///
    f32_a f32_al f35_a f35_b hist_0 hrtany inv_0 mpage mupdate mptype
occombdu occombtr para0 parity podu potr ///
    p_cyodu p_cyotr ses smoy tumor_1 tumor_2 tumor_3 f37 smdu f10 f11
eore pore p_cyore ep_cyre ep_ctre occombre ///
    eetr eedu eere le_vgtr le_potr

```

```

/* Check PNR */

```

```

/* Before */
pnrcheck pnr, listw di

```

```
replace pnr="194107049043" if lopnr==11069
replace pnr="193205068947" if lopnr==11107
replace pnr="194312284609" if lopnr==11126
replace pnr="191910261666" if lopnr==11195
replace pnr="193401073329" if lopnr==11279
save"C:\usr\rino\projects\ovarian_new\data\tomas\tomas.dta", replace

/* after */
pnrcheck pnr, listw di

log close
```