Subject: Re: merging male and female hiv data to couples data (DRC)
Posted by Bridgette-DHS on Wed, 20 Jul 2016 14:25:30 GMT

View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Every time you do a merge, a diagnostic variable called "_merge" is constructed by Stata. You can use it to decide what cases to keep. For example "keep if _merge==3" will reduce the merged file to just those cases which are in both of the original files. Then you need to either rename that variable, for example with rename _merge _merge_IR_AR" or drop it with "drop _merge". If you do not do that, subsequent merges will be blocked, as you have found.

In this survey, there are 18,257 people in the AR file. Only 183, that is, 1%, were HIV positive. There are so few HIV positive cases that virtually no analysis is possible. If, say you want to look at HIV discordance among the couples, there will be very few such couples.

Also the HIV testing was not done on all men and women. You can look up the selection procedures. There are 18,827 women in the IR file, 8,656 men in the MR file, 18,257 men+women in the AR file, and 4,486 couples in the CR file. This means there will be many couples for whom you cannot get the HIV status of both the man and the woman.

Having said that, I will paste below the possible steps for this merge. At the end I produce a variable that gives the combination of the man's and the woman's HIV statuses.

I did a report on this topic in 2013 (https://www.dhsprogram.com/pubs/pdf/AS35/AS35.pdf). You may want to look at that.

* Merge CR and AR files for CongoDR

set more off
use e:\DHS\DHS_data\AR_files\CDAR61FL.dta, clear
rename hivclust v001
rename hivnumb v002
rename hivline v003
rename hiv03 vhiv03
rename hiv05 vhiv05
keep v*

sort v001 v002 v003
save e:\DHS\scratch\CDAR61_vtemp.dta, replace

use e:\DHS\DHS_data\AR_files\CDAR61FL.dta, clear
rename hivclust mv001
rename hivnumb mv002
rename hivline mv003
rename hiv03 mvhiv03
rename hiv05 mvhiv05

```
keep mv*

sort mv001 mv002 mv003
save e:\DHS\scratch\CDAR61_mvtemp.dta, replace

use e:\DHS\DHS_data\CR_files\CDCR61FL.dta, clear
sort v001 v002 v003
merge v001 v002 v003 using e:\DHS\scratch\CDAR61_vtemp.dta
rename _merge _merge_woman

sort mv001 mv002 mv003
merge mv001 mv002 mv003 using e:\DHS\scratch\CDAR61_mvtemp.dta
rename _merge _merge_man

tab _merge*,m

* then decide what combinations of the two _merge variables to keep

gen couple_hiv_status=.
replace couple_hiv_status=1 if vhiv03==0 & mvhiv03==0
replace couple_hiv_status=2 if vhiv03==0 & mvhiv03==1
replace couple_hiv_status=3 if vhiv03==1 & mvhiv03==0
replace couple_hiv_status=4 if vhiv03==1 & mvhiv03==1

label define status 1 "W- M-" 2 "W- M+" 3 "W+ M-" 4"W+ M+"
label values couple_hiv_status status

tab couple_hiv_status,m
```