
Subject: Cannot replicate IPV ever for India DHS 2005-06

Posted by [dgodha](#) on Tue, 07 Mar 2017 15:51:20 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hello,

I have been trying to replicate IPV ever for India DHS 2005-06 (NFHS III) but while I am able to get the population size correctly, the prevalence estimate does not match for the 15-19 age group (Table 15.9 on page 509). There is a similar issue raised for Nepal but trying that solution has not worked here. I have used the following commands using Stata 13.1:

```
/*Generating weights*/
gen dwt=d005/1000000
svyset [pw=dwt], psu(s021) strata(v025)

*Any intimate partner violence
egen ipv=rsum (d104 d106 d107 d108)

gen violence=0 if ipv==0
replace violence=1 if ipv>0 & ipv~=.

svy :tab violence if v013==1 & v501>=1 & v501<=5 & v044==1
```

I will really appreciate if someone can point me the reason for the discrepancy and how to get the exact match.

Thanks

D. Godha

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by [Liz-DHS](#) on Fri, 10 Mar 2017 15:13:31 GMT

[View Forum Message](#) <> [Reply to Message](#)

Dear User,

Thank you for posting to the forum. I have referred your post to one of our experts. In the meantime, please take a look at these series of YouTube videos. If you find your answer in one of these videos, please do let us know. Thank you!

[https:// www.youtube.com/playlist?list=PLagqLv-gqpTMU3avlnBDodTWCazUR y4CT](https://www.youtube.com/playlist?list=PLagqLv-gqpTMU3avlnBDodTWCazUR y4CT)

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by [Bridgette-DHS](#) on Fri, 17 Mar 2017 14:48:51 GMT

[View Forum Message](#) <> [Reply to Message](#)

Following is a response from Senior DHS Specialist, Kerry MacQuarrie:

The problem is in the code below when you sum d104-d108 to create the variable ipv. Variables

d104, d106, d107, and d108 have some "9" codes which are getting picked up along with the "1"s. You only want to count the "1"s. This means that women who did not report that form of violence are getting swept into the yeses and her estimate on the variable "violence" is higher than reported in the Final Report. Hope this helps!

```
/*Generating weights*/  
gen dwt=d005/1000000  
svyset [pw=dwt], psu(s021) strata(v025)
```

```
*Any intimate partner violence  
egen ipv=rsum (d104 d106 d107 d108)
```

```
gen violence=0 if ipv==0  
replace violence=1 if ipv>0 & ipv~.
```

```
svy :tab violence if v013==1 & v501>=1 & v501<=5 & v044==1
```

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by [dgodha](#) on Fri, 17 Mar 2017 17:43:26 GMT

[View Forum Message](#) <> [Reply to Message](#)

Dear Kerry,

I appreciate your help here.

Just to clarify further on your reasoning, this is not the case with India DHS 2005-06. If there had been any '9's, then I would have coded differently and included only the '1's in my counting. Further proof that I am not dealing with any '9's is that the maximum count my variable 'ipv' has is 4 (range: 0-4).

I will request your assistance to get to the root of the problem.

Thanks

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by [Bridgette-DHS](#) on Mon, 20 Mar 2017 11:00:49 GMT

[View Forum Message](#) <> [Reply to Message](#)

Following is a response from Senior DHS Specialist, Kerry MacQuarrie:

There are two issues; my first response only addressed one. First, I am looking at the NFHS-3 dataset (IAIR52fl.dta) and I see 9's. My % for each form of IPV separately (emotional, physical,

sexual) match the table, as shown below. Perhaps you are not working with the same data file I am (IAIR52fl.dta) or you have already recoded the 9's to 0's.

```
. tab1 d104 d016 d107 d108 [iw=dwt]
```

ever any
emotional

violence	Freq.	Percent	Cum.
----------	-------	---------	------

no	56,073.474	84.12	84.12
yes	10,532.383	15.80	99.92
9	51.7936259	0.08	100.00

Total	66,657.651	100.00	
-------	------------	--------	--

experienced
any less
severe

violence (d105a-d)	Freq.	Percent	Cum.
-----------------------	-------	---------	------

no	48,021	69.11	69.11
yes (d105a-d)	21,404	30.80	99.92
9	59	0.08	100.00

Total	69,484	100.00	
-------	--------	--------	--

experienced
any severe
violence

(d105e-g)	Freq.	Percent	Cum.
-----------	-------	---------	------

no	62,455	89.88	89.88
yes (d105e-g)	6,972	10.03	99.92
9	57	0.08	100.00

Total	69,484	100.00	
-------	--------	--------	--

experienced
any sexual
violence

(d105h-i)	Freq.	Percent	Cum.
-----------	-------	---------	------

no	59,955.996	89.95	89.95
yes (d105h-i)	6,640.2286	9.96	99.91

9 61.42587743 0.09 100.00

Total 66,657.651 100.00

gen PV=.

replace PV=0 if (d106!=. & d106!=1) | (d107!=. & d107!=1)

replace PV=1 if d106==1 | d107==1

. ta PV [iw=dwt]

PV	Freq.	Percent	Cum.
----	-------	---------	------

0	43,293.625	64.95	64.95
---	------------	-------	-------

1	23,364.026	35.05	100.00
---	------------	-------	--------

Total 66,657.651 100.00

However, there is a second problem and that is in creating the composite ipv and violence variables; Yours estimate prevalence of any form at 31.6% instead of 39.7%. The following code works to produce estimates that match the table, for each form of IPV separately and for the combinations of physical + sexual and physical + sexual + emotional. The code is not as efficient as yours was, but it is correct.

use "IAIR52FL.dta", clear

/*Generating weights*/

gen dwt=d005/1000000

svyset [pw=dwt], psu(s021) strata(v025)

*Recode 9's

recode d104(9=0),g(d104_r)

recode d106(9=0),g(d106_r)

recode d107(9=0),g(d107_r)

recode d108(9=0),g(d108_r)

*Physical violence

g PV=d106_r

replace PV=1 if d107_r==1

*Physical or sexual violence

g PVSV=PV

replace PVSV=1 if d108_r==1

*Any intimate partner violence

g ipv=PVSV

replace ipv=1 if d104_r==1

tab1 d104_r PV d108_r PVSV ipv [iw=dwt]

svy :tab ipv if v013==1 & v501>=1 & v501<=5 & v044==1

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by [dgodha](#) on Mon, 20 Mar 2017 14:16:59 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hello Kerry,

I appreciate your looking into the matter further.

You were right about the dataset. Apparently I had the earlier file- IAIR50FL. So, I re-downloaded the latest one.

But coming to the second part, there is no difference between our commands (they are asking the software to do the same thing) or the output. And I ran both on the IAIR52FL file and I am still getting the same answers- so back to square one. Our population sizes (for "Any violence": physical or sexual or emotional) match with that in the report-4643 but the percentage comes out exactly 33.72. Does this mean 33.72% for "Any violence" is correct?

Your command:

```
use "IAIR52FL.dta", clear
/*Generating weights*/
gen dwt=d005/1000000
svyset [pw=dwt], psu(s021) strata(v025)
```

*Recode 9's

```
recode d104(9=0),g(d104_r)
recode d106(9=0),g(d106_r)
recode d107(9=0),g(d107_r)
recode d108(9=0),g(d108_r)
```

*Physical violence

```
g PV=d106_r
replace PV=1 if d107_r==1
```

*Physical or sexual violence

```
g PVS=PVS
replace PVS=1 if d108_r==1
```

*Any intimate partner violence

```
g ipv=PVS
replace ipv=1 if d104_r==1
svy :tab ipv if v013==1 & v501>=1 & v501<=5 & v044==1
```

My command:

```
use "IAIR52FL.dta", clear
/*Generating weights*/
gen dwt=d005/1000000
svyset [pw=dwt], psu(s021) strata(v025)
```

*Recode 9's

```
recode d104(9=0),g(d104_r)
recode d106(9=0),g(d106_r)
recode d107(9=0),g(d107_r)
recode d108(9=0),g(d108_r)
```

```
egen ipv1=rsum (d104_r d106_r d107_r d108_r)
tab ipv1
```

```
gen violence=0 if ipv1==0
replace violence=1 if ipv1>0 & ipv1~=.
svy :tab violence if v013==1 & v501>=1 & v501<=5 & v044==1
```

Thanks in advance
DG

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06
Posted by [Bridgette-DHS](#) on Mon, 20 Mar 2017 15:14:37 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from Senior DHS Specialist, Kerry MacQuarrie:

My code below does produce a different estimate than yours and one that matches the final report (39.7% ever experienced any IPV). Your population is higher (83,703) on your composite variable than what it should be (66,658). Your code, using the egen command, can be retained and will produce the correct estimate/population if you modify the code for creating the composite "violence" variable as follows:

```
gen violence=0 if ipv1==0
replace violence=1 if ipv1>0
replace violence=. if d005==. | v502==0
```

This removes from the 0 category and places into missing both women who did not complete the domestic violence module AND women who did complete the domestic violence module but did NOT answer questions d104-d108 on spousal abuse because they are never married. This produces an estimate of 39.7% experiencing any IPV from a population of 66,658 as shown in the final report.

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06
Posted by [dgodha](#) on Tue, 21 Mar 2017 05:47:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello Kerry,

Let me reiterate my original question:

"I have been trying to replicate IPV ever for India DHS 2005-06 (NFHS III) but while I am able to get the population size correctly, the prevalence estimate does not match for the 15-19 age group (Table 15.9 on page 509)."

Now, I have applied your correction to my command but I am still getting 37.2% for IPV ever for the 15-19 year age group but the weighted population is a match at 4643.

My commands:

use "IAIR52FL.dta", clear

```
/*Generating weights*/
gen dwt=d005/1000000
svyset [pw=dwt], psu(s021) strata(v025)
```

```
*Recode 9's
recode d104(9=0),g(d104_r)
recode d106(9=0),g(d106_r)
recode d107(9=0),g(d107_r)
recode d108(9=0),g(d108_r)
```

```
egen ipv1=rsum (d104_r d106_r d107_r d108_r)
tab ipv1
```

```
gen violence=0 if ipv1==0
replace violence=1 if ipv1>0 & ipv1~=.
replace violence=. if d005==. | v502==0
```

```
svy :tab violence if v013==1 & v501>=1 & v501<=5 & v044==1
```

As I had mentioned in my earlier email, I have tried your commands too and I get the same output for the 15-19 year age group.

I will look forward to your feedback.

Thanks
DG

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06
Posted by [dgodha](#) on Wed, 22 Mar 2017 05:36:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

Sorry, that was 33.72% not 37.2%.

I look forward to your response.

Thanks
DG

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06
Posted by [Bridgette-DHS](#) on Wed, 22 Mar 2017 11:37:11 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from Senior DHS Specialist, Kerry MacQuarrie:

The correct numbers are produced when missings (9's) are correctly clubbed with "no". The correct numbers to match should be as below.

File Attachments

1) [dv.jpg](#), downloaded 1773 times

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06
Posted by [dgodha](#) on Wed, 22 Mar 2017 11:45:27 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks, that's all I needed to know. The prevalence of "Emotional, sexual, or physical violence" in the 15-19 year age group is 33.72% as opposed to 33.9% shown in Table 15.9 on page 509 in the chapter "Domestic Violence" of NFHS-3 Report (India DHS 2005-06).

I appreciate your patience.
Thanks again

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06
Posted by [yemsokha](#) on Thu, 01 Nov 2018 21:09:11 GMT
[View Forum Message](#) <> [Reply to Message](#)

Dear senior,

Would you mind explain me why we need to use this code:

*Recode 9's

recode d104(9=0),g(d104_r)

recode d106(9=0),g(d106_r)

recode d107(9=0),g(d107_r)

recode d108(9=0),g(d108_r)

Bests,

sokha

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by [Liz-DHS](#) on Fri, 02 Nov 2018 13:29:27 GMT

[View Forum Message](#) <> [Reply to Message](#)

Dear User, that code is used to deal with the missing values. The 9s indicate missing data. Thank you!

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by [yemsokha](#) on Fri, 02 Nov 2018 15:22:37 GMT

[View Forum Message](#) <> [Reply to Message](#)

Dear seniors,

I am working with Cambodia DHS 2014, dataset khir73.dta.

My objective is to compare the IPV violence between rural and urban setting.

IPV were combined of 3 kinds of violences(IPV=sexual violence+emotional violence+physical violence).

Would you mind tell me that whether my combination is correct or not?

```
gen wt=d005/1000000
```

```
svyset [pw=wt], psu(v021) strata (v023)
```

```
svy: tab v013 d101a if v044==1 & v502!=0, row count
```

```
tab d105h
```

```
recode d105h 2/3=1
```

```
tab d105h
```

```
tab d105i
```

```
recode d105i 2/3=1
```

```
tab d105i
```

```
gen sv=d105i+d105h
```

```
tab sv
```

```
tab d105k
```

```
recode d105k 2/3=1
```

```
tab d105k
```

```
gen sv1=sv+d105k
```

```
tab sv1
```

```
recode sv1 2/3=1
```

```
tab sv1
```

```
****sv1=sexual violence
```

```
tab d105b
```

```
recode d105b 2/4=1
```

```
tab d105b
```

```
tab d105a
```

```
recode d105a 2/3=1
```

```

tab d105a
gen pv=d105a+d105b
tab pv
recode pv 2=1
tab pv
tab d105c
recode d105c 2/4=1
tab d105c
gen pv1=pv+d105c
tab pv1
recode pv1 2=1
tab pv1
tab d105d
recode d105d 2/3=1
tab d105d
gen pv2=pv1+d105d
tab pv2
recode pv2 2=1
tab pv2
tab d105e
recode d105e 2/3=1
tab d105e
gen pv3=pv2+d105e
tab pv3
recode pv3 2=1
tab pv3
tab d105f
recode d105f 2/3=1
tab d105f
gen pv3=pv2+d105f
gen pv4=pv3+d105f
tab pv4
recode pv4 2=1
tab pv4
tab d103c
recode d103c 2/4=1
tab d103c
tab d103a
recode d103a 2/4=1
tab d103a
gen ev=d103c+d103a
tab ev
recode ev 2=1
tab ev
tab d129
recode d129 2=1
tab d129
gen ev1=ev+d129

```

```

tab ev1
recode ev1 2=1
tab ev1
tab d103b
recode d103b 2/4=1
tab d103b
gen ev2=ev1+d103b
tab ev2
recode ev2 2=1
tab ev2
drop ev2
tab d101a
recode d101a 8=.
tab d101a
gen ev2=ev1+d101a
tab ev2
recode ev2 2=1
tab ev2
tab d101b
recode d101b 8=.
tab d101b
gen ev3=ev2+d101b
tab ev3
recode ev3 2=1
tab ev3
drop v3a08a v3a08b v3a08c v3a08d v3a08e v3a08f v3a08g v3a08h v3a08i v3a08j v3a08k v3a08l
drop v3a08n v3a08o v3a08q v3a08r v3a08s v3a08t v3a08u v3a08v
drop v3a08x v3a08z m55z_1 m55z_2 m55z_3 m55z_4 m55z_5 m55z_6 v463z
drop v465 v467f h15f_1 h15f_2 h15f_3 h15f_4 h15f_5 h15f_6 h37y_1 h37y_2 h37y_3 h37y_4
h37y_5 h37y_6 v603 v604 v633b v779
tab d101d
recode d101d 8=.
tab d101d
gen ev4=ev3+d101d
tab ev4
recode ev4 2=1
tab ev4
drop v822 v850b s826f
tab d101e
recode d101e 8=.
tab d101e
gen ev5=ev4+d101e
tab ev5
recode ev5 2=1
tab ev5
tab v739
tab d101f

```

```
recode d101f 8=.
tab d101f
gen ev6=ev5+d101f
tab ev6
recode ev6 2=1
tab ev6
tab d101c
recode d101c 8=.
tab d101c
gen ev7=ev6+d101c
tab ev7
recode ev7 2=1
tab ev7
tab v467b
tab s817a
recode s817a 8=.
tab s817a
gen ev8=ev7+s817a
tab ev8
recode s817a 2=1
tab ev8
recode ev8 2=1
tab ev8
gen vaw=ev8+sv1+pv4
tab vaw
recode vaw 2/3=1
tab vaw
Bests,
```

sokha

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by [Liz-DHS](#) on Fri, 09 Nov 2018 19:38:38 GMT

[View Forum Message](#) <> [Reply to Message](#)

Dear User, A response from senior analyst, Dr. Kerry MacQuarrie:

Quote:

Hello,

I am reluctant to assess an IPV variable as "correct" or "incorrect" since it is really up to the researcher's discretion, depending on your analytic purpose. Are you trying to construct the spousal violence variable(s) as The DHS Program constructs them in, for example, Table 20.9 Forms of Spousal Violence (p 268) in the Cambodia DHS 2014 Final Report? If so, which numbers are you trying to match?

The code at the end of this message should work for most surveys to produce estimates of any spousal violence (emotional, physical, or sexual) in the past 12 months--the last results column of

Table 20.9. The code would be a bit different to produce the first results column (EVER any spousal violence).

Here a couple of resources that I find helpful when I am trying to match DHS survey report tables.

1. The Guide to DHS Statistics, available here (see Ch. 17):

[https://www.dhsprogram.com/publications/publication-DHSG1-DH](https://www.dhsprogram.com/publications/publication-DHSG1-DHS-Questionnaires-and-Manuals.cfm)
[S-Questionnaires-and-Manuals.cfm](https://www.dhsprogram.com/publications/publication-DHSG1-DHS-Questionnaires-and-Manuals.cfm)

2. DHS tutorial videos on Matching DHS Final Report Tables Part 1

(https://www.youtube.com/watch?v=tjuaPV4eC_E) and Part 2

(<https://www.youtube.com/watch?v=TovRq9TpbXU>)

In summary, the steps to matching a table are to check if you have:

1. Correct data file unit of analysis
2. Correct denominator population at risk
3. Correct variables find and understand your variables
4. Correct recoding special values
5. Correct weights
6. Correct tabulation row vs. column percent

I hope these resources help you.

Regards,
Kerry

Stata code to produce estimates of spousal violence (emotional, physical, sexual, & any of these 3) in past 12 months

```
use [CC]IR[VV]FL.DTA, clear /*CC=country code; VV=version number*/  
gen dwt = d005/1000000
```

```
**Create composite variable "experienced ANY spousal violence in last 12 months"
```

```
/*NOTES:
```

1. Restricts sample to women selected for the domestic violence module.
2. Restricts sample to ever-married women (currently married refers to current husband; formerly married refers to last husband)
3. Domestic violence weight should be used for all analyses with this variable.

```
*/
```

```
*Emotional violence variables
```

```
recode d103a (1/2=1 "Yes") (0 3 =0 "No"),g(humil)  
recode d103b (1/2=1 "Yes") (0 3 =0 "No"),g(threat)  
recode d103c (1/2=1 "Yes") (0 3 =0 "No"),g(insult)
```

```
egen ev12=rowtotal(humil threat insult) if humil!=.
```

```
recode ev12 (1/3=1)
```

```
lab def yesno 0 "No" 1 "Yes"
```

```
lab val ev12 yesno
```

```
lab var ev12 "Experienced any emotional violence in last 12 months"
```

*Physical violence variables

```
recode d105a (1/2=1 "Yes") (0 3 =0 "No"),g(push)
recode d105b (1/2=1 "Yes") (0 3 =0 "No"),g(slap)
recode d105c (1/2=1 "Yes") (0 3 =0 "No"),g(punch)
recode d105d (1/2=1 "Yes") (0 3 =0 "No"),g(kick)
recode d105e (1/2=1 "Yes") (0 3 =0 "No"),g(burn)
recode d105f (1/2=1 "Yes") (0 3 =0 "No"),g(knife)
recode d105j (1/2=1 "Yes") (0 3 =0 "No"),g(twist)
```

```
egen pv12=rowtotal(push slap punch kick burn knife twist) if humil!=.
```

```
    recode pv12 (1/max=1)
```

```
    lab val pv12 yesno
```

```
    lab var pv12 "Experienced any physical violence in last 12 months"
```

*Sexual violence variables

```
recode d105h (1/2=1 "Yes") (0 3 =0 "No"),g(sex)
recode d105i (1/2=1 "Yes") (0 3 =0 "No"),g(sexact)
recode d105k (1/2=1 "Yes") (0 3 =0 "No"),g(psexact)
```

```
egen sv12=rowtotal(sex sexact psexact) if humil!=.
```

```
    recode sv12 (1/3=1)
```

```
    lab val sv12 yesno
```

```
    lab var sv12 "Experienced any sexual violence in last 12 months"
```

*Any spousal violence composite variable

```
egen dv12=rowtotal(ev12 pv12 sv12) if humil!=.
```

```
    recode dv12 (1/3=1)
```

```
    lab val dv12 yesno
```

```
    lab var dv12 "Experienced any spousal violence in last 12 months"
```

```
ta dv12
```

```
ta dv12 [iw=dwt]
```

Kerry LD MacQuarrie, PhD

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by anikhpg42@gmail.com on Thu, 21 Nov 2019 11:35:39 GMT

[View Forum Message](#) <> [Reply to Message](#)

Dear Seniors and Dr. Kerry MacQuarrie,

Thank you for your Stata code to produce estimates of spousal violence (emotional, physical, sexual, & any of these 3) in past 12 months.

I have used it for NGDHS_IR_2018 dataset and completely matched.

But, would you kindly provide the Stata code for estimating EVER spousal violence (emotional, physical, sexual, & any of these 3) in lieu of 'past 12 months'?

Thank you very much.

Subject: Re: Cannot replicate IPV ever for India DHS 2005-06

Posted by [Liz-DHS](#) on Fri, 22 Nov 2019 19:13:19 GMT

[View Forum Message](#) <> [Reply to Message](#)

Dear User, a response from Dr. Kerry Macquarrie:

Quote:

There are two ways to code EVER spousal violence.

The first way is to alter the code you have for spousal violence in the last 12 months. That code considers 1 and 2 (often and sometimes in the last 12 months, respectively) as yes. You would want to consider 3 (yes, but not in last 12 months) as yes as well. So any place that the code reads 1/2=1, it should read 1/3 =1, as in:

```
recode d103a (1/2=1 "Yes") (0 3 =0 "No"),g(humil)
```

becomes

```
recode d103a (1/3=1 "Yes") (0=0 "No"),g(humil)
```

The other way is to use the existing variables in the dataset. Experienced any emotional violence is d104 and experienced any sexual violence is d108. To get experienced any physical violence, you will need to combine d106 and d107:

```
gen phys_viol=.
```

```
replace phys_viol=1 if d106==1 | d107==1
```

```
replace phys_viol=0 if d106==0 & d107==0
```

Kerry LD MacQuarrie, PhD
