Subject: Plasma blood glucose in 2011 and 2017/2018
Posted by MiFoo on Fri, 29 Jan 2021 00:33:39 GMT
View Forum Message <> Reply to Message

Hi everyone,

I am working with the PR files from 2011 and 2017 in order to analyze diabetes prevalence in Bangladesh. These are my questions:

1) I tried to replicate some of the tables in the final reports. My results are very similar but not equivalent. I guess the number of included observations is the most likely reason for these deviations. For instance, my dataset contains 13147 valid observations (0=no, 1=1) for the variable sb325, while the final report from 2017 states 12,946 respondents (p232). Similarly, my dataset contains more observations for the variables measuring plasma blood glucose (sb335b in 2017 and sh284a in 2011) than the tables in the reports (e.g. p235 in 2017). Are some respondents excluded from the analysis? Why? I am using the household weight for my calculations (hv005). Please correct me if this is not the correct approach (even though this should not influence the raw number of respondents).

2) According to the codebook, the variable sb335b in 2017 is defined as "Plasma blood glucose (mmol/dl) - 1 implied deci". Might this be an error in the codebook? Given the distribution of values it seems like the unit of measurement is 10*mmol/l.

Thanks!

Subject: Re: Plasma blood glucose in 2011 and 2017/2018
Posted by Trevor-DHS on Mon, 08 Feb 2021 17:45:30 GMT
View Forum Message <> Reply to Message

1) It looks to me that you are not using the weights. When I use the following, I get roughly the right number (it is differing by 3 cases which I think is due to some missing data):
. tab sb325 [iw=hv005/1000000] if sb325 == 1 | sb325 == 0

```
   glucose |
  measured |
    before |     Freq.    Percent      Cum.
------------+-----------------------------------
        no | 9,828.9921     75.94     75.94
       yes |3,114.39667     24.06    100.00
------------+-----------------------------------
     Total | 12,943.389    100.00
```

Yes, some cases from the dataset are not used, specifically those that refused to answer the questions about glucose or refused to be tested.

For the table 13.5.3 on page 235 I match the number of cases using the following code:
tab sb335b [iw=hv005/1000000] if sb302==2 & sb303c >= 35 & sb335b < 900,m

This results in 3437 cases matching the table. Here I have included women age 35+ who have valid measures for glucose.

2) The data are recorded with 1 implied decimal place. That is a value of 125 should have one implied decimal place and is thus 12.5, or a value of 52 should have one implied decimal place and is thus 5.2. The label is a bit confusing and is as you say 10*mmol/L.

Subject: Re: Plasma blood glucose in 2011 and 2017/2018
Posted by MiFoo on Wed, 17 Feb 2021 21:55:40 GMT
View Forum Message <> Reply to Message

Thank you very much for your explanations!
You are right: I did not use the weights when calculating the number of nonmissing values. I thought I read somewhere that the number of respondents reported in DHS tables would be the "raw" number but obviously I am mistaken.

Subject: Re: Plasma blood glucose in 2011 and 2017/2018
Posted by Trevor-DHS on Thu, 18 Feb 2021 01:34:25 GMT
View Forum Message <> Reply to Message

The numbers in DHS tables are always weighted numbers except when the table explicitly says they are unweighted.