
Subject: Sibling mortality in maternal IR files
Posted by [bun_2019fall](#) on Wed, 09 Mar 2022 16:43:55 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi DHS program.

I have recently encountered a question about a dataset: DRIR4BFL (Dominican Republic, 2002 Standard DHS). I found that not all women sampled have information on their siblings (around half of women sampled have this information). I am not sure if this information is collected randomly. That is, I wonder if the women who are missing sibship information are random sample.

- I checked the documentation of that survey, and I do not think the documentation specifically specifies the missing. Rather, the information I saw in the dataset does not seem to correspond with the documentation. However, the sample size of women having sibship information I found is exactly the same as a publication using the same data: Obermeyer Z, Rajaratnam JK, Park CH, Gakidou E, Hogan MC, et al. (2010) Measuring Adult Mortality Using Sibling Survival: A New Analytical Method and New Results for 44 Countries, 1974-2006. PLoS Med 7(4): e1000260. doi:10.1371/journal.pmed.1000260

- I also tried to retrieve information from the DR 2002 DHS Report; however, that report is in Spanish.

I wonder if anyone would kindly direct me to places / people where I can resolve this issue? Any help or directions are very much appreciated!

Thank you!

Subject: Re: Sibling mortality in maternal IR files
Posted by [Bridgette-DHS](#) on Thu, 10 Mar 2022 12:35:23 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is response from DHS Research & Data Analysis Director, Tom Pullum:

If you enter "tab midx_01,m" in the IR file, there are 11,280 cases with the value 1 and 12,104 cases with a dot. When you see this pattern for one of the modules, it means that there was subsampling. Half of the households would have been selected for the this module. The subsampling was probably done to reduce data collection costs. If you look at hv117 and hv118 in the PR file, you will also see that men were also subsampled, with a smaller fraction. Usually when there is subsampling you can find a variable in the PR or IR file with labels such as "selected for sibling module" and "selected for men's survey". Unfortunately, I can't find such variables in the files for this survey. They may be in there but I don't see them. In many of the older surveys the subsampling indicator was dropped during data processing.

Subsampling is not left up to the interviewers. There would have been some selection, almost certainly of households, using a random procedure, during the household listing before the cluster

was entered by the interviewer teams. The 12,104 consists mainly of women whose household was not selected, but also includes some women who had no eligible siblings. That's probably the only reason why the module was NA for more than half of the women.

Subject: Re: Sibling mortality in maternal IR files
Posted by [bun_2019fall](#) on Fri, 11 Mar 2022 02:43:35 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thank you for much for the detailed response and clarification here. I really appreciate it!

I likewise could not find specific information regarding the selection of this subsample of the DR 2022 survey. While another survey (Kenya, 2014; also ~50% women did not have sibship information) also only has a subsample, but that survey's documentation records the reason in which that full version was only administered to a subsample (although it is unclear how the subsample was selected).

Thank you for the clarification and the information again!
