
Subject: DHS representativeness by cohorts
Posted by [DHS user](#) on Wed, 22 Jan 2020 13:02:44 GMT
[View Forum Message](#) <> [Reply to Message](#)

I'm working with kid recode analyzing the inequality of opportunity in health (IO) (using height-for-age) and comparing the differences in IO by cohorts (0-1 year, 1-2, 2-3, 3-4, 4-5).

I know that generally the sample is representative at national, regional and urban-level, that it is necessary take into account sampling design, etc. I read also Tom Pullum said in the forum that "the surveys are representative of the population living in household, regardless of age, sex, etc".

But... my question is the following: is the sample in the kid recode representative also by cohorts of children? I don't read anything about it and I would appreciate so much your answer.

Subject: Re: DHS representativeness by cohorts
Posted by [Bridgette-DHS](#) on Wed, 22 Jan 2020 13:05:06 GMT
[View Forum Message](#) <> [Reply to Message](#)

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

Within selected households, apart from surveys of ever-married women, all women age 15-49 are selected for the women's interview and those who are de facto residents are retained for the IR file. Some of the information about children age 0-4 is collected in the household interview, for all children in the household, and some other information is collected during the women's interview, regarding her children. The KR file includes all children age 0-4 whose mothers are in the household survey, and if the child was also in the household, relevant data from that survey, such as the anthropometry data, is attached to the child.

The children identified in the household survey are representative of all children in the household population. The children identified in the women's survey are representative of all children who are in the household population AND have a surviving mother. The KR and BR files do include children who are not living with their mother or have died, but if the child is not living with the mother, then some information in the files is routinely dropped from tabulations. For example, the mother is asked, for children age 0-4, whether the child had diarrhea in the past two weeks, and if yes, whether it was taken for medical treatment. However, if the mother is not living with the child, the tabulations drop the child, because there is good evidence that mothers tend to under report illness for children who are not living with them. This may introduce some bias.

Thus I would consider the children in the survey to be representative of all children, but there are some potential biases if the mother has died or is not living with the child.

If the number of children in some combination of characteristics is small, there will of course be an increase in the standard errors of estimates based on those children. However, apart from what I just mentioned, the estimates will be unbiased.

Just one other comment, although this may be obvious. The term "cohort", as in "birth cohort", applies to children born about the same time. The health outcomes for a cohort only refer to the survivors. Some children, of course, have died. That may be because they were severely undernourished, were not immunized, were not provided with a bednet and had malaria, etc. Death is the worst possible outcome, and when we look only at the survivors, we risk missing the bigger picture. Nevertheless, we do tend to ignore the children who died, because we don't have the right data regarding their nutrition, immunizations, etc., and therefore can't make inferences about WHY they died.
