
Subject: Corrected age from individual file - India 2019
Posted by [Thiago Melo](#) on Tue, 27 Sep 2022 18:10:46 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello,

While working with some child indicators in the PR dataset of India 2019, we noticed some differences in the ages provided by the following variables:

hv105 age of household members, used in indicators such as "Percentage of de jure children under age 5 whose births are registered with civil authorities and who had a birth certificate"
hml16 corrected age from individual file, used in indicators such as "Percentage of children under age 5 who slept the night before the survey under an insecticide treated net (ITN)"

We have some instances where the same person has 0 years in one variable and 100 years in other (see Figure 1 attached). This leads to differences in the indicators' denominators. We also checked India 2015 and Gambia 2019 for comparison and the differences were not so dramatic (Figures 2 and 3, respectively).

Why are there such big differences?

We would appreciate some guidance.

Thank you in advance,

Thiago Melo
International Center for Equity in Health | Pelotas

File Attachments

- 1) [Figure 1.png](#), downloaded 201 times
 - 2) [Figure 2.png](#), downloaded 205 times
 - 3) [Figure 3.png](#), downloaded 201 times
-

Subject: Re: Corrected age from individual file - India 2019
Posted by [Janet-DHS](#) on Wed, 28 Sep 2022 18:48:30 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS staff member Tom Pullum:

Age in the household survey (hv105) is provided by the household respondent, as part of the listing of household members. There can be differences from the ages given for women in the women's survey (v012) or men in the men's survey (mv012) because those ages involve more probing and are asked of the specific woman or man. For children, hv105 may not match up with b8 in the birth history or the other appearances of age and birthdate as they appear in the various files, including the PR file.

The year, month, and day of birth for a specific child are the same wherever they appear in the data files. Age on the day of data collection is exactly consistent with birthdate. The only potential discrepancy is that the age in years that is calculated from birthdate and day of interview may not be consistent with hv105, as you observe. The calculated age (or, for adults, v012 and mv012) should always be given priority over hv105.

Some analysis of potential deficiencies in hv105 are described in chapter 2 of this methodological report: <https://www.dhsprogram.com/pubs/pdf/MR26/MR26.pdf>. Sometimes the differences between hv105 and b8 (or other appearances of age for children) are substantial. If you see both 0 and 100 for the same person, there is clearly a data entry error.

Actually it's a great strength of DHS that the data files include both hv105 and the calculated age. You can use discrepancies to assess possible data quality problems, and compare different interview teams or different surveys. But even the calculated age, although it is better than hv105, is subject to displacement and other kinds of measurement error in many contexts if the respondents have no reason to keep track of birthdates.
