



STEP 3.2: Create the sampling frame and analytic variable data sets



What is the purpose of this step?

In this step you will:

- Create the child-level sampling frame data set (specified in Step 2).
- Create the child-level analytic variable data set (specified in Step 2).
- If Applicable: Deliver the data sets to your survey vendor.
- If Applicable: Have the survey vendor pull your PHDS sample.



Guidelines and Issues to Consider

- Create the child-level sampling frame data set and analytic variable data set.

Step 2.2 provided a detailed description of the guidelines that should be used in identifying and selecting the sample.

To review, you (or your hired vendor) will do the following:

- 1) Identify eligible children for sampling.
- 2) Determine the minimum starting sample for each unit of analysis.
- 3) Stratify the starting sample by age and determine the minimum starting sample for each age group in each unit of analysis.

Once these steps have been completed, you are ready to randomly select the starting sample for the PHDS.

Additionally, for each child and unit of analysis represented in the starting sample you will identify the analytic information not based on the survey, as described in **Step 2.3**.

- ☑ If Applicable: Provide survey vendor with a sampling and analytic data set data dictionary (**Appendix 7**).

If you are providing a data set to a survey vendor to conduct the sampling, outline the specific variables you want to include in the child-level sampling and analytic data sets (e.g., utilization, enrollment, etc.), and/or obtain a data dictionary from your survey vendor before the PHDS sample is pulled. This ensures that all variables that need to be linked to the survey data set will be available to you. Once the sample is pulled, you may not be able to go back to get additional variables. In particular, be sure your vendor makes clear which stratified sampling groups each child belongs to (e.g., age, continuous enrollment, age group, and any state-specific sampling strategies).

Tip from the Field

You want to use the most updated information possible for sampling because a key eligibility criterion is whether the child had a visit or not. Children have seven well-child visits in their first year of life alone. Therefore, you want to make sure that your sampling is done as close as possible to when the survey is administered. Past users of the PHDS have conducted a “dry run” of the sampling to ensure that their sampling data specifications work and then have re-run the sampling at the latest time possible.