STEP 2.2: Plan and confirm the feasibility of your PHDS sampling strategy

What is the purpose of this step?

The purpose of this step is to ensure you will identify a starting sample that will allow you to analyze the PHDS data in a way that meets your project goals.

In this step you will:

☑ Verify all units of analysis for which you would like to construct PHDS measures.

☑ Identify eligible children for sampling.

☑ Specify the minimum completed and starting sample sizes needed for each unit of analysis.

☑ Specify age-stratifications required for each specific unit of analysis.

☑ Finalize and implement your sampling strategy.

Guidelines and Issues to Consider

☑ Verify your unit(s) of analysis

This may seem like an obvious step, and you may have already done this while conceptualizing your project. However, it is critical that you are clear about your unit(s) of analysis. In other words, what entities, areas, or groups of children are you measuring? Each of these units of analysis has different specifications for sampling, so you need to decide upon them first.

Units of analysis past users of the PHDS have sampled for include:

1. **Health System**

   For example: Across the entire pediatrics department in a health plan.

2. **Office or Provider-Group**

   For example: An office located at a specific location that is comprised of multiple providers.

3. **Individual Health Care Providers**

   For example: Dr. Jones and Dr. Smith

4. **Specific Populations of Children**

   For example: Children within specific race-ethnicity groups. Children who reside in specific locations (i.e., rural vs. urban).

Secondly, you need to ensure that there are valid data to allow you to analyze the desired unit of analysis. **Table 2.1** located on the next page provides highlights of common units of analysis for the PHDS and issues to consider when sampling for these specific groups. The bottom line is to think about what unit of analysis is most relevant to your priority audiences for the PHDS findings. Health care systems vary across markets, so who is accountable and who will use the information will differ depending on your health care system.

Table 2.1 Units of Analysis and Sampling Issues to Consider

Potential Unit of Analysis	Issues to Consider at the Time of Sampling
Health System	➲ Variations in the "microsystems" within the health system in how preventive and developmental health care is provided**:** Consider the characteristics within the "microsystems" that influence how well-child care is provided. Consider the various groups within a health system that may want to analyze the findings.
Office or Provider Groups	➲ Number of providers: When doing office-level sampling, it is imperative to consider the number of providers and the (full-time employees) of the providers in each office. Bigger offices will need a bigger starting sample size than smaller offices. ➲ Provider team: Think about the health care provider team that will be measured. How is well-child care provided? Who gives the care that is measured in the PHDS? Do the nurse and physician divide up the well-child visit? The more people who provide care measured in the PHDS, the more variation there will be, and the greater the sample size will need to be.
Individual Health Care Providers	➲ Provider team (see above) ➲ Provider-level variables: What variables will you use to identify the provider to whom the child should be assigned at the time of sampling? There are two options: 1) Provider with whom the child is enrolled or "paneled" as a primary care provider. 2) Provider with whom the child had the most well-child visits in the last 12 months or since birth. **Tip from the Field:** Although you may want to analyze the data at the health care provider level, you may not have valid information at that level. Specifically, the CAHMI team has found that while many systems note the provider with whom the child is enrolled, this provider is not necessarily the person that the parent is most likely to think about when responding to the survey. This could be due to a variety of factors: A) The provider variable is based on the claims/bills database (this allows for one centralized billing code for a medical group) or B) The child may receive well-child care from providers with whom they are not enrolled (e.g. they see a provider in the same office, therefore the claim is still paid, etc.). To address this issue, **CAHMI recommends** that systems use available enrollment and utilization information for sampling and then ask parents/survey respondents to indicate their child's personal doctor or nurse in the survey and use this data for provider-level reporting.

Table 2.1 Units of Analysis and Sampling Issues to Consider (Continued)

Potential Unit of Analysis	Issues to Consider at the Time of Sampling
Specific Populations of Children	➲ Child-level variables: What variables will you use to identify specific children? Are these variables valid? **Tip from the Field:** Many health systems have variables related to a child's race/ethnicity in their data systems, but these variables are not reliable (e.g., they are not consistently used or they are only valid for some children). To address this issue, **CAHMI recommends** that systems use available enrollment and utilization information for sampling and then ask parents/survey respondents to indicate their child's race/ethnicity and use this data for reporting purposes.

☑ Identify **eligible children** for sampling

Sampling is the process used to identify children whose parents will be asked to complete the PHDS. There are five criteria that a child must meet in order to be eligible for inclusion in the PHDS sample:

1. **Age criterion**: Select children 3–45 months of age (allows for time lag in sampling administration).

   Purpose: The PHDS measures care recommended in the first three years life. This includes care provided through the three-year-old well-child visit.

2. **Enrollment criterion**: Select children continuously enrolled in the health system for 12 months or since birth.

   Purpose: The PHDS is a measure of health care quality. You want to include in the sample children who have been in the system for the time period referenced in the survey.

3. **Well-child visit criterion**: Select children who have had one or more well-child visits (as defined by the HEDIS well-child visit specifications) in the last 12 months or since birth.

   Purpose: The PHDS is a measure of health care quality. The care provided in the PHDS should have been provided during well-child appointments. You want to include in the sample children who have received well-child care in the time period that is referenced in the survey.

4. **One-child per family/target child criterion**: Randomly select only one eligible child per family.

   Purpose: You want the PHDS administration process to be as family-centered as possible. **CAHMI recommends** that only one child be sampled for the survey, as many families could have more than one eligible child and may be overwhelmed by multiple surveys.

5**. Give survey in language spoken in the home** (if available)**:** Select families that speak the language in which the survey is administered.

   Purpose: The current version of the PHDS is available in English and Spanish and the ProPHDS is available only in English at this time (translation will occur in Fall 2006).

Tip from the Field

Before going on to the next step, **CAHMI recommends** that you examine the number of children that meet the eligibility criterion described above for each unit of analysis. For example, if you are sampling for individual providers, we recommend that you first examine how many children are eligible for each of the providers. This will raise issues early on in the process that may otherwise arise once the sampling strategy is implemented.

Additional PHDS Resources: Keep in mind that if you are planning to administer the survey in the pediatric office (not by mail) or via the telephone, a different sampling methodology will need to be used (see guide for *In-Office Administration of the PHDS, Reduced Item Version* or the *PHDS-PLUS Implementation Guidelines* listed in the resources section).

☑ Specify the minimum completed and starting sample sizes needed for each unit of analysis

Now that you have identified eligible children, you need to specify the following:

1) The minimum number of completed surveys that you will need for each unit of analysis**.**
2) You can then determine the starting sample size needed, taking into account the following:

a) The response rate you think you will be able to achieve,

b) The number of surveys that will not reach the parent for completion due to bad addresses, and

c) The data error rate for the specific until of analysis.

Figure 2.1: Determining starting sample size

$$\text{Minimum sample} = \frac{\text{target completed surveys}}{\text{response rate–bad address rate) * (1–data error rate)}}$$

The sampling strategy that you implement is dependent on how you will be using the results. For example, if you plan on using the results to compare health care providers, then you will need more completed surveys than if you were using the results to examine the quality of preventive care at the population level.

Tip from the Field

If you are planning multiple uses for your results, choose the sampling strategy with the largest minimum sample required.

Table 2.2 provides recommended sampling strategies based on different units of analysis. Definitions of each of the variables in this table are located on the next page.

Table 2.2: Determining the Starting Sample Size Required for Each Unit of Analysis

	Comparison of individual providers [a]	Comparison of offices or provider groups	Health-system level reporting	Comparison of health plans [e]
Target number completed surveys	30 per health care provider [b]	30 per health care provider in each office	100[d]	250 per health plan
Estimated data error rate	1%	1%	1%	1%
Estimated response rate	40%	40%	40%	40%
Bad address rate	Depends on the setting	Depends on the setting	Depends on the setting	Depends on the setting
Minimum starting sample, assuming no bad addresses [c]	78 per health care provider	78 per health care provider in each office	253	632 per health plan

[a] Although a smaller sample could be drawn if you are not planning on using the results for comparison, we recommend that you assume comparisons will be made if you are reporting results at the provider or health plan levels. If 30 surveys are not feasible, the minimum number CAHMI recommends per provider is 15. See Table 2.1 for other issues to consider in provider-level sampling. Lastly, one of the PHDS measures (follow-up for children at risk) is only calculated for a portion of children (approximately 25% of the sample). Therefore, if this is a primary measure to be used in comparisons, then the sample size should be adjusted accordingly.

[b] Providers who are very consistent in the care they provide across patients will need fewer surveys, as compared to providers who target certain discussions to certain patients. Secondly, if the provider and nurse each provide components of the well-child visit, then more surveys may be needed as the provision of care by two individuals increases the level of variation in this communication-dependent measure.

[c] CAHMI recommends that each sample contain members enrolled in the same type of health insurance coverage. Therefore, different samples should be drawn if you wish to assess quality of care for Medicaid beneficiaries and commercial enrollees.

[d] As is described in Table 2.1, the more providers there are, the more variation there is. Therefore, CAHMI recommends that you base the sample size on the number of providers. An alternate approach is to base the sample size on the number of FTE in each office.

[e] This is the minimum number of surveys recommended. However, to date, all of the Medicaid agencies and recent health plans that have implemented the PHDS have set their completed survey goal at N=2000. This sample size has allowed the state to do a number of analyses that met their strategic and political goals, and allowed stratified analysis for specific groups of children and program and policy areas.

Definition of Terms in Table 2.2

Targeted Number of Completed Surveys: The minimum number of completed surveys necessary for analysis. A completed survey is defined as a survey in which at least 80 percent of the items were answered; it will be discussed in greater detail later in this section.

Estimated Response Rate: The percentage of parents who responded to the survey. You can never know for sure what the response rate for your survey will be. However, you can estimate this rate based on your own previous survey experience. If you do not have previous experience, we recommend using 40 percent. This represents a conservative estimate for a response rate based on field-testing and the implementation of similar surveys. It should ensure that you have enough completed surveys for analysis. Field test and previous implementation results of the PHDS have yielded response rates from 20 to 70 percent. Any response rate estimate that you have from previous survey experience in your area should be substituted for the estimated response rate when determining the minimum sample size. Many factors that can influence the response rate of your survey, and suggestions will be provided throughout this section to help you to maximize your response rate.

Bad Address Rate: The rate of addresses in your database that will be incorrect. As is noted later in the chapter, Address Service Correction should be part of your survey administration. However, there still will be a number of addresses that will be incorrect. The rate of bad addresses has varied significantly across past users of the PHDS (2%–38%). CAHMI recommends that you examine other surveys used within the health system to determine an estimated bad address rate.

Estimated Data Error Rate: The rate of data errors that you expect within your sample or sampling frame. Data errors are incorrect or bad contact information, enrollment information, eligibility information, or any other type of information necessary for the administration of the survey. You may not know what the data error rate is for your sampling frame; however, you are likely to find some data errors. We recommend using a rate of 1 to 2 percent if you do not know your rate. If you do know the data error rate, this number can be substituted in the chart above and will increase or decrease your minimum sample size.

Minimum Starting Sample: The minimum number of children who should be sampled for the administration of the survey given the intended use of the results.

EXAMPLE 2.2: Determining Minimum Sample Size

The Health Plan A chose to administer the PHDS across the entire system. The plan primarily contracts directly with 10 medical groups comprised of 25 individual providers and would like to use the PHDS results primarily for quality improvement at the system-, office- and provider level. However, they will also be publishing the results in a consumer guide. Two years ago, the plan administered the CAHPS and had a 52 percent response rate and a 3 percent bad address rate. They conducted an audit of their provider records just last year and expect their data error rate to be less than 0.05 percent.

Minimum starting sample size=
25 (*# of providers*)* 30 (*# of completed surveys per provider*)
(.52 (*response rate*)–.03 (*bad address rate*)) * [1–(.05) (*data error rate*)]

Minimum starting sample = 1611

☑ Specify age stratifications required for each specific unit of analysis

The last step in identifying the starting sample of children whose parents will be sent the PHDS survey is to stratify the sample for three age groups of children.

- Children **3–9.99** months old at the time of survey administration
- Children **10–18.99** months old at the time of survey administration
- Children **19–45.99** months old at the time of survey administration

This stratification is to ensure that sufficient samples are obtained for the three groups listed above. The reason you want sufficient samples for each of these age groups is because the PHDS items focused on anticipatory guidance and parental education are different for each of these groups.

At the time of sampling, it is important to specify the date when the surveys will be sent out to the parent and to conduct the age-stratification based on how old the child will be when the parent receives the first mailing of the survey.

There are two options for stratifying the sample by age that have been used:

Option 1:

- One-fourth of the starting sample is children **3–9.99** months old
- One-fourth of the starting sample is children **10–18.99** months old
- One-half of the starting sample is children **19–45.99** months old

Option 2:

- Examine the proportion of eligible children (the children who met the five eligibility criterion described earlier) in each of the groups in your health system for the units of analysis of interest and base the stratification on your own population.

It is important that each sample is stratified for each unit of analysis. For example, if you are sampling for 10 offices. The sample **for each office should be stratified by age**.

Tip from the Field

In order to reduce burden and administrative time, users of the PHDS for office- and provider-level analysis have stratified the sample at the office level only. However, if you are using the PHDS for incentive-based payments, then the starting sample should be stratified for each unit of analysis examined.

Important Note for Users of the ProPHDS:

If you are using the ProPHDS, you will administer three, age-specific versions of the survey that map to these three age-specific groups. The age of the child for the starting sample needs to be adjusted to allow for the time of survey administration. CAHMI recommends that you assume that the survey administration will take two months. Therefore, the starting sample for the ProPHDS should be stratified by the following age groups.

- One-fourth of the starting sample is children **3–7.99** months old
- One-fourth of the starting sample is children **10–16.99** months old
- One-half of the starting sample is children **19–43.99** months old

☑ Finalize and implement sampling strategy

Once you have:

1) Identified eligible children for sampling
2) Determined the minimum starting sample for each unit of analysis
3) Stratified the starting sample by age

You are then ready to randomly identify the number of children in each group in each of the units of analysis.

Tip from the Field

Again, we recommend confirming the feasibility of obtaining all needed data before finalizing your plan. It is not uncommon for data elements or contact information needed to administer the PHDS to be lacking for key subgroups of children who you would like to include in your sampling.