



## STEP 4.1: Monitor survey administration

### ➔ What is the purpose of this step?

The purpose of this step is to closely track the quality of your survey administration process, and to monitor and troubleshoot any problems with securing an optimal response rate and accurate data entry.

#### In this step you will:

- Receive and review weekly tracking reports from your survey vendor.
- Assess and troubleshoot problems with response rates.



### Guidelines and Issues to Consider

- Specify and obtain weekly survey tracking report.

Survey responses should be tracked throughout the administration process. At a minimum, a variable indicating the survey disposition should be maintained by your survey vendor and reported on a weekly basis throughout the administration process. Values for this variable can include:

- *Survey received* – Survey was sent back.
- *Complete* – At least 80 percent of the survey items parents should answer were completed.
- *Incomplete* – Less than 80 percent of the survey items were completed.
- *Refused* – Parent refused to complete survey, sent it back blank, OR called the toll-free line and asked to be removed from list.
- *Bad Address*
- *Ineligible* – Parent returned the survey saying that their child was not in the age group, has not had a well-child visit, or child is no longer enrolled in the health system.
- *Deceased* – Child has passed away.
- *Language barrier* – Respondent cannot read and understand the survey.

Calculate interim response rates and troubleshoot response problems.

Your *raw response rate* is simply the number of questionnaires returned divided by the number of children in your starting sample. However, this rate is often misleading and not representative of the true response rate. Therefore, an *adjusted response rate* should be calculated on an interim basis throughout survey administration to track whether you are meeting response rate targets.

For the adjusted response rate, both the numerator (number of respondents) and the denominator (number sampled) are adjusted based on certain factors:

- First, the numerator (number of respondents) is adjusted from the total number of questionnaires returned to the number of completed questionnaires returned. A survey is defined as “completed” if the parent answered at least 80 percent of the items that they should have answered. You should not include items that were skipped appropriately in your count. You can adjust this algorithm if you want to allow responses to be considered even if less than 80 percent of items are completed.

Second, the denominator (number sampled) is adjusted to include only those children who are eligible for the survey, for whom you have a valid address, and whose parents do not have a language barrier. The formula to adjust your response rate is:

$$\text{Response Rate} = \frac{\text{Number of 80\% or more completed surveys}}{\text{Number of children sampled} - (\text{deceased} + \text{ineligible} + \text{language Barrier} + \text{bad address})}$$

If your response rate falls short of projections, you may need to troubleshoot with your survey vendor and consider adding and/or altering administration steps. For instance, you may decide to add an additional phone call, send a postcard, alter your introductory message left on voicemail systems, or introduce a small incentive for participation in the survey.



## STEP 4.2: Clean and code your PHDS data set

### ➡ What is the purpose of this step?

The purpose of this step is to obtain an accurate data set from your survey vendor. This step includes consistent and accurate cleaning and coding of the data set in preparation for constructing PHDS quality measures and the analytic variables needed to report your PHDS project findings.

#### In this step you will:

- Specify data files to be created.
- Obtain and check interim data sets for accurate data labeling and entry.
- Prepare data files for analysis.



### Guidelines and Issues to Consider

While data preparation is often considered part of the analysis, this task is included in the data collection section because it may be completed by the vendor while administering the survey. Additional data preparation and cleaning steps are described in Step 5.

- Specify data files to be created.

If you are using a vendor to administer the survey, the vendor should submit a data file that contains the following:

1. Coded responses for all PHDS items, including blank, do not know, refused to answer, and item skipped.
2. Survey disposition, such as if the survey was completed and the reasons for incomplete surveys (see variables noted in Step 4.1).
3. Other descriptive variables identified and collected for the starting sample (e.g., related to enrollment, health care utilization, etc.) that were identified in Steps 2 and 3.
4. Administrative data used for generating the sampling frame.
5. Age of child in months.
6. Any supplemental data linked prior to the removal of identifying information used for survey administration.
7. A data dictionary for the file. An example data dictionary for the ProPHDS survey data file can be found in **Appendix 9**.

You also may choose to have your vendor conduct some initial data preparations, such as:

- ⇒ Verifying ineligible cases
- ⇒ Checking for duplicate data records
- ⇒ Running frequencies on all variables to check for values that are out of range
- ⇒ Identifying problems with skip patterns

If errors are found, you should have the vendor verify them with the original surveys to ensure that the errors stem from the respondent and not from the administration process. Once these are identified, you will need to make decisions on how they will be handled for the analysis. Refer to Section 5 for more detail on analyzing the results.

- ☑ Obtain and check "test" and interim data sets for accuracy of data entry and survey administration.

The vendor administering your survey (either internally or externally) should provide you with a test and interim PHDS survey data sets according to a predetermined schedule.

**CAHMI recommends** that you ask your survey vendor to send a test data set that is based on a handful of mock completed surveys. This data set will test the data entry processes and ensure that the data set you receive matches the data dictionary your vendor is using.

#### **Tips from the Field**

- Always label data variables.
- Update your data dictionary with any changes made to data labels or response codes. Good documentation is essential!
- Create a backup of your data set in case of emergency. Also, create temporary and permanent data sets wisely. Think about what you would need to do if you lost the data.
- Always keep a copy of your original data set.

When you receive the test data set you should make sure that your vendor is using the data variable labels agreed upon and that responses to survey items are assigned the agreed-upon values (e.g., 1 = "no"; 2 = "yes", etc.). If errors are found, request that they be corrected immediately.

CAHMI recommends that you request at least two interim data sets. The first should include the first 100 surveys entered and the second should include half of your expected final completed survey data set (e.g., if your final complete survey goal is N=2,000, then the second interim data set should be N=1,000). These interim data sets allow you to develop the syntax that you will use to clean and analyze your PHDS data. Therefore, when the final data set is received, you will have already done a significant amount of preparatory work.

Preparing the PHDS data files for analysis.

Valid PHDS findings require careful preparation of your data prior to starting your analysis. The following are necessary steps to prepare the data for analysis. They do not necessarily need to be conducted in the order presented.

**Data Prep Step #1:** Verify survey completeness.

You should receive the data from the vendor for all of the interviews conducted. However, for your analysis you should limit the data to surveys with at least 80 percent of the items completed.

**Data Prep Step #2:** Check for ineligible cases.

Make sure parents who responded have children who meet the sampling criteria for age and continuous enrollment. (a) Run a frequency on the age variable from the survey responses. Here you should ensure that the age the parent reports in the survey and the age-specific section of the PHDS that the parent completes match the age of child that you have in your administrative data files. Use the parent report as the "gold standard" and exclude cases where the child was erroneously included in the sampling frame. (b) Remove records where the child was found not to be in the health plan, provider, or unit you are sampling.

**Data Prep Step #3:** Check for duplicate data records.

Make sure every record has a unique identifier.

**Data Prep Step #4:** Check for out-of-range values.

Run frequencies on all of your variables to check for out-of-range values or odd-looking distributions. At this point, you may not be able to go back and correct the data error. If the error is random and affects only a few cases, then you may want to exclude those cases. However, if the error seems to be systematic and affects a large number of responses, it may be worth finding the source of the error and correcting it.

**Data Step #5:** Identify problems with skip patterns.

Run frequencies and cross-tabulations to verify that skip patterns were followed correctly. If errors seem random and affect only a small number of records (less than 2%), assume the item stem (the question instructing the respondent to go to a different question) is accurate and then correct the response for the incorrect skip. Systematic errors or problems with a significant number of cases should be verified.

If you want to be absolutely certain that skip patterns were followed, you can require that only the children of parents who responded appropriately to the filter question are included when you create the new variables.

**Data Prep Step #6:** Assign missing values.

Missing values should be recoded in some way so that you know not to include them in the analyses. You should designate missing values in the data set in a way that ensures they are omitted when calculating measures. Also, recode the response options of "refused" to a missing value. Examine the number of "I don't know" responses that you get. If this total percentage is less than 2 percent, then you should recode them as missing values.



## STEP 4.3: Integrate updated child enrollment and utilization information

### ➔ What is the purpose of this step?

The purpose of this step is to make sure that data for each child are as current as possible when you calculate your PHDS quality indicators and create reports of your findings for key audiences.

#### In this step you will:

- Identify variables that may need to be updated for each sampled child.
- Obtain and merge updated data variables into your analytic variable data set.



### Guidelines and Issues to Consider

- Identify variables that may need to be updated for each sampled child.

As discussed in Steps 2.2, 2.3, and 3.2, many analytic variables are derived from enrollment and utilization data sets and are valuable for analyzing and reporting PHDS findings.

A majority of the PHDS asks the parent to respond about care provided "in the last 12 months." Therefore, you want the variables in the analytic data file to accurately describe the 12 months prior to the date when the parent responded to the survey.

Examples of common variables that need to be updated from the time of sampling include:

- ⇒ Variables related to health care utilization: There is often a lag between when a child has a visit and when it shows up in the data systems. You want to wait to update your analytic file until the data systems are updated and correct for when the parent completed the survey.

*For example: If there is a three-month lag before a visit shows up in your data systems, and the survey was administered in September, then you want to wait until December to update your analytic data file with the information about health care utilization.*

⇒ Variables related to enrollment: Examples of these variables include the primary care provider with whom the child is enrolled.

- ☑ Obtain and merge updated data variables into your analytic variable data set.

Using the child-level unique identifier specified in Step 2.3, you merge the updated data variables into your analytic variable data set. This analytic variable data set will be what you use in Step 5 of this manual.



## **STEP 4.4: Identify and obtain additional descriptive information about the health system to inform analysis**

### **➔ What is the purpose of this step?**

The purpose of this step is to collect additional descriptive information about the health system to inform your analysis. This information should address the specific processes, systems, and resources that relate to the provision of preventive and developmental health care. This information can be used to identify which processes and systems are correlated with higher and poorer quality of care, and then inform policy and quality improvement decisions.

#### **In this step you will:**

- Identify health system characteristics related to the provision of preventive and developmental health care.
- Collect this descriptive information for each unit of analysis.



### **Guidelines and Issues to Consider**

- Identify health system characteristics related to the provision of preventive and developmental health care.

This important step will enhance the value and usefulness of the PHDS data in guiding future policy and improvement efforts. Review the PHDS survey items again and think about specific processes, systems, and resources in place that are related to the topics measured in the PHDS.

Below is a description of tools that can help you gather related descriptive information:

## 1. Office Systems Inventory

**Appendix 10** provides an example of the Office Systems Inventory (OSI) developed through the Healthy Development Collaborative.<sup>1</sup>

The OSI collects descriptive information about specific office systems related to:

- meeting parents' informational needs and addressing their concerns about their child's learning, development, and behavior;
- identifying children at risk through the use of structured developmental and psychosocial assessments and screening at appropriate visits;
- providing strong links to community resources for families who need or want them; and
- promoting optimal parent/child relationships.

The OSI can be completed by the office manager or another office staff member who has the knowledge needed to complete the tool.

The PHDS can then be analyzed by the office-specific OSI scores, and systems yielding higher quality of care scores can be identified.

### **Tip from the Field**

Where possible, you should collect information on the OSI for each office included in your starting sample. For example, if you are using the PHDS to examine quality of care in 10 pediatric offices, then you should collect the OSI for each office.

---

<sup>1</sup> The Healthy Development Collaborative was a Commonwealth Fund–supported initiative designed to help primary care practices in Vermont and North Carolina engage families in a partnership to promote positive developmental outcomes for the families' children through the development of improved office systems. The OSI is part of the *Practical Guide for Healthy Development*, a set of materials and tools designed and tested in the Healthy Development Learning Collaborative, a 12-month quality improvement initiative. The OSI is also available at [http://www.cmf.org/usr\\_doc/Office\\_Systems\\_Inventory.pdf](http://www.cmf.org/usr_doc/Office_Systems_Inventory.pdf).

## 2. Posters and Handouts in the Office

Parents can receive information about the topics measured in the PHDS through posters and handouts located in the office of their child's health care provider. These resources may answer parents' questions and meet their informational needs without involving a discussion with their child's health care provider.

Given that the PHDS allows the parent to indicate whether they had their informational needs met, it can be useful to catalog the types of information available to the parent and analyze the PHDS findings using this descriptive information.

**Appendix 11** provides an example of a worksheet that can be used to catalog posters and handouts. The worksheet lists all of the topics that are included in the PHDS and has a column to indicate whether there is a related poster or handout that is provided in the office. It is important to catalog only those educational resources that are readily available to the parent and do not require a discussion with the office staff, as such discussions are included in the PHDS.

The PHDS can then be analyzed by the poster and handout scores, and educational resources correlated with higher quality of care scores can be identified.

## 3. Electronic Medical Records

Some health systems utilize electronic medical records (EMRs). Most EMR programs have standardized forms or templates that can be used by the provider for well-child visits. And most EMR programs allow users to develop standard parent education templates (often called an "After-Visit Summary") that are handed out by the health care provider. The worksheet provided in **Appendix 11** can also be used to catalog standard EMR text and handouts.

- Collect this descriptive information for each unit of analysis.

In Step 2.2 you specified the units of analysis for sampling and analysis. It is important to collect as much descriptive information for each unit of analysis as possible, as it will enhance the useability of your PHDS findings and will help you to identify possible reasons for high and lower performance within each unit of analysis.



## STEP 4.5: Weight your data set to represent your target population

### ➡ What is the purpose of this step?

The purpose of this step is to ensure that your final PHDS data set is representative of the desired sample of children on which you would like to report at the end of your project. Weighting is a standard component of most survey projects, especially if you oversampled for certain population subgroups (e.g., racial groups) or if there are response biases (e.g., some population subgroups are systematically less likely to respond to your survey).

#### In this step you will:

- Clarify your desired population of children to ensure that your findings are representative and determine whether weighted data are required.
- Develop a weighting methodology and construct sampling weights to use during your data analysis.



### Guidelines and Issues to Consider

- Determine whether weighting is required.

The goals of the sampling and survey administration strategy outlined in this manual are:

- 1) To obtain sufficient samples of completed surveys for the three age groups that correspond to the age-specific sections of the PHDS survey (3–9 months, 10–18 months, and 19–48 months).
- 2) To obtain specific units of analysis to focus your measurement effort.

You will need to weight your data if you want to be able to create reports that describe findings across your health system for all children.

Specifically, you will need to weight your data if one or more of the following is true about your PHDS project:

- ⇒ You over-sampled for specific groups.
- ⇒ You used age-stratifications that are not representative of your population.
- ⇒ The responding population is significantly different than your sampled population.

☑ Develop a weighting methodology.

Your survey department or vendor should be able to construct valid weights once you specify the population of children on which you want to report PHDS findings (e.g., all children enrolled in your health system who had at least one-well child visit).

Below are issues and tips to consider as your weighting method is developed.

1. Clarify the baseline population you will use to calculate the weights.

- ❖ Since the children eligible for the PHDS include ONLY those who met the eligibility criteria described in Step 2.2 (e.g., continuous enrollment and visit requirement; parents spoke the languages in which the survey was administered; and families had valid address information), you can only weight data to represent this group of children.
- ❖ A primary purpose of weighting is to ensure the PHDS findings accurately describe care for a group of children. It is important to clarify the group of children for which you want the report to provide generalizable findings.

For example:

- ⇒ Do you want the report to describe care provided across an entire health system? If so, then the baseline population is all eligible children enrolled in the health system.
- ⇒ Do you want the report to provide office-specific descriptive information? If so, office-specific weights will need to be created based on the eligible population of children in each of the pediatric offices.

2. Specify the factors you will use to calculate weights.

- ❖ At a minimum, you will need to weight your data according to age if your general population of children is not distributed in the way you stratified the sample in Step 2.2.
- ❖ To make your PHDS findings representative, you will need to determine the actual age distribution of the population of eligible children, according to the three groups specified in Step 2.2, and weight the data accordingly.
- ❖ Other variables you may need to include in your weighting methodology include:
  - 1) Variables you used for over-sampling such as race, geographic location, and type of health care provider or health plan.
  - 2) Variables for populations that are under- or over-represented in your completed sample due to a known response bias.