

## PBS Substudy Design Memo No. 2012-04

---

DATE: July 03, 2012  
TO: PBS Study Centers  
FROM: NCS Field Support Team  
SUBJECT: Sampling women within selected providers

**Purpose:** The purpose of this memo is to provide guidance to PBS Study Centers on choosing the optimal and most appropriate subsampling method(s) for provider locations in order to select women for recruitment.

### Methods for Sampling Women

Each provider location will have a sampling rate assigned to it based on the location's measure of size (MoS), in this case the number of first prenatal care visits practitioners in the location have. PBS sites must adhere to this sampling rate when selecting women to recruit. However, determining how to best implement the sampling rate will depend in part on the preferences and arrangements made with the provider location. Please see *Appendix 1 - Schemes for Sampling Women within Selected Providers* for more details on how the sampling rate will be calculated by the methods explained below.

In addition to the sampling rate, there are four major factors to consider when determining the logistics for selecting sampled women for a given provider location:

- **Eligibility Screening:** When is the Eligibility Screener applied (real time in the office vs. retrospective)?
- **Patient List:** Who builds the patient list (NCS staff or provider locations staff)?
- **Prescreening:** Can or will the provider prescreen on any eligibility criteria?
- **Identifying Patients:** How will patients be identified? There are several options:
  - The provider will release the PII to NCS directly.
  - The provider will ask sampled patients for permission to release PII.
  - The provider will refer patients to contact the NCS themselves.
  - The provider will allow NCS to approach women as they come or go from appointments in the location.

More specific guidance on these logistical considerations and how they might drive the choice of patient selection method appears at the end of this memo, after a series of example scenarios.

Once these logistical considerations are determined, SCs can review them in conjunction with the sampling rate for that provider location to determine the best methods of subsampling women and recruiting them into the NCS. There are four methods of subsampling women within a provider location. The four methods are:

1. Systematic subsampling only (every Nth woman on a list);
2. Temporal subsampling only (every Xth week in the data collection period);
3. Temporal and systematic subsampling (every Nth woman in every Xth week); and
4. Two layers of temporal subsampling (subset of days in every Xth week).

Below, we provide a more detailed explanation of each of the four methods, how each method would be employed, and some example scenarios of how they might work logistically in the provider office. These scenarios also shed some light on how the four logistical factors help determine the ideal selection method for a given location.

For each scenario, we also show an example of a provider listing table. The listing table is where the list of women from the provider location is captured. The listing table combines data that could come from four different sources into a single summary. These sources are:

- Data obtained directly from the provider;
- Data values implied by sampling methods/pre-screening arrangements;
- Data values created by the Study Center; and
- Data obtained from patient during the eligibility screener.

As shown in the tables below, certain pieces of information (e.g., patient’s name, pregnancy status) will be obtained or created differently depending on the logistical arrangements with the provider location. The tables use the following color-code scheme to display how the information is obtained in each scenario:

	Information from provider
	Implied by prescreening
	SC creates information
	Obtained during eligibility screener

*Appendix 1: Methods for Sampling Women within Selected Providers* provides further details to illustrate how the 4 sampling methods could potentially be applied within provider offices under some hypothetical sampling rates. For each sampled provider location, such a chart will be generated by Westat based on that provider locations particular sampling rate to display the sampling method options and parameters. Each of the subsampling methods are described below.

## **Method 1: Systematic Subsampling Only**

In this method, the location will build a list of patients during the entire 4-month recruitment period. The SC will systematically sample women from the list (select every Nth woman); Westat will provide instructions for how to execute the sampling of women based on the subsampling rate.

### *Scenarios*

1. A small location with a subsampling rate of 0.5, or in other words selecting one in every two women.

This office can allow NCS staff to access their records. The location has an electronic records system and will allow SC staff to come in once a week to build a list of patients who came in that week. The system will allow the SC to determine the woman's age, whether or not she is pregnant, and whether it was the first visit at this location for this pregnancy. The SC will sample from the list and follow-up with selected patients to screen for eligibility.

An example patient listing table appears below in Table 1.1.

The location supplied: the date of visit, whether the patient was age eligible, pregnant, and whether it was her first prenatal visit at this location. The SC staff will only extract the PII for those patients who are potentially (not prescreened as ineligible) eligible.

The SC will create an ID for each line on the list, and then filter the list down to the potentially eligible patients (lines #1, #2, #5, #7, #9, #10 and #13). They will then subsample every other patient from the pool of potentially eligible patients.

In some locations, the provider staff will introduce the Study to patients in order to obtain permission to release their PII to the SC. The outcome of this introduction is captured in the MDES variable Provider Intro Outcome. In this case, for the patients who passed the prescreening, it will be set to "5 Provider can release PII" because the location will not be introducing the Study to the patient. For the patients who did not pass the prescreening, the Provider Intro Outcome is set to "-7 - Not Applicable."

Person ID	Date of Visit	First Prenatal Visit At Location? (Yes/No)	Pregnant? (Yes/No)	Age Eligible? (Yes/No)	In PSU? (Yes/No)	Prescreen Status	Provider Intro Outcome	Sampled Person (Yes/No)	Person First Name	Person Last Name	Phone Number
1	6/1/2012	Y	Y	Y		In	5 = Provider can release PII	Y	Jane	Doe	555-555-1212
2	6/1/2012	Y	Y	Y		In	5 = Provider can release PII	N	Teresa	Garcia	555-555-1313
3	6/1/2012	N	Y	Y		Out	-7 (Not Applicable)	n/a			
4	6/1/2012	N	N	Y		Out	-7 (Not Applicable)	n/a			
5	6/4/2012	Y	Y	Y		In	5 = Provider can release PII	Y	Beth	Jones	555-555-1414
6	6/4/2012	N	Y	Y		Out	-7 (Not Applicable)	n/a			
7	6/4/2012	Y	Y	Y		In	5 = Provider can release PII	N	Mary	Smith	555-555-1515
8	6/4/2012	N	N	N		Out	-7 (Not Applicable)	n/a			
9	6/4/2012	Y	Y	Y		In	5 = Provider can release PII	Y	Anna	Carpenter	555-555-1616
10	6/5/2012	Y	Y	Y		In	5 = Provider can release PII	N	Rachel	Doyle	555-555-1717
11	6/5/2012	N	Y	Y		Out	-7 (Not Applicable)	n/a			
12	6/6/2012	N	Y	Y		Out	-7 (Not Applicable)	n/a			

	2										
13	6/6/2012	Y	Y	Y		In	5 = Provider can release PII	Y	Jae	Lee	555-555-1919

**Table 1.1: Systematic Subsampling Only (Method 1), Scenario 1**

Key:

Information from provider	Implied by prescreening	SC creates information	Obtained during eligibility screener
---------------------------	-------------------------	------------------------	--------------------------------------

*Method 1 Scenarios (continued)*

2. A location has a subsampling rate of 0.5, or in other words selecting one in every two women.

This location has agreed to generate a list of all patients with non-identifying ID numbers that they can match to their records, but they will not prescreen the list at all. The SC will sample from the list and inform the location which IDs were sampled. The location will obtain permission to release PII, either by mailing a form, calling the patient, or at the next visit to the location.

A sample patient listing table appears below in Table 1.2. The location only initially supplies an ID number and date of visit. After the SC performs the sampling, the location will supply the PII for women who have agreed to be contacted.

The Provider Intro Outcome will be set to 2 or 3, reflecting whether or not the patient agreed to release her PII. In this example, Patients 1, 3, 5 and 7 were sampled, and 1, 3, and 7 agreed to let the provider release their PII. The values for first prenatal visit, pregnant, and age eligibility, and residency in the county will be captured in the eligibility screener in this scenario, so they are not displayed in this example table.

Person ID	Date of Visit	First Prenatal Visit? (Yes/No)	Pregnant? (Yes/No)	Age Eligible? (Yes/No)	In County? (Yes/No)	Prescreen Status	Provider Intro Outcome	Sampled Person (Yes/No)	Person First Name	Person Last Name	Phone Number
1	6/1/2012					No Prescreen Performed	2 = Female Agreed to SC Contacting Her	Y	Jane	Doe	555- 555-1212
2	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	N			
3	6/1/2012					No Prescreen Performed	2 = Female Agreed to SC Contacting Her	Y	Beth	Jones	555- 555-1414
4	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	N			
5	6/1/2012					No Prescreen Performed	3 = Female Refused to Allow SC to Contact Her	Y	Refused	Refused	Refused
6	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	N			
7	6/1/2012					No Prescreen Performed	2 = Female Agreed to SC Contacting Her	Y	Jae	Lee	555-555-1919
8	6/1/201					No Prescreen	-7 (Not Applicable)	N			

	2					Performed					
--	---	--	--	--	--	-----------	--	--	--	--	--

**Table 1.2: Systematic Subsampling Only (Method 1), Scenario 2**

**Key:**

Information from provider	Implied by prescreening	SC creates information	Obtained during eligibility screener
---------------------------	-------------------------	------------------------	--------------------------------------



## Method 2: Temporal Sampling Only

In this method, entire weeks would be sampled (1 in X weeks). The value of X is chosen such that systematic sampling of women from the list is not needed; all women who visit the provider location during the selected time period would be selected. Also, the value of X should be chosen so that at least two full work weeks are included during the recruitment period. Given the four month (16 week) field period for the PBS recruitment effort, Method 2 is only able to be used in locations with a sampling rate of .125 or higher (2 weeks out of 16). Ideally, the weeks will be randomly selected (not determined by the location) and non-consecutive. Study Centers will work with Westat to randomly select weeks within the recruitment period.

### *Scenarios*

1. This is a medium-sized location with a subsampling rate of 0 .125 (1 in 8).

Working with Westat, the 7<sup>th</sup> and 15<sup>th</sup> week of the field period are selected. Every woman visiting the office during those two weeks will be eligible. The NCS staff will be given a desk in the waiting area. This provider location will prescreen patients on pregnancy, first prenatal visit at this location, and age. Provider staff will describe the Study to pregnant adult women at the end of their first prenatal appointments and ask women to stop by the NCS desk on the way out. At the end of each day, the data collector will need to ask the provider location for two pieces of information: the total number of patients who were seen by a practitioner at that location that day, and the number of women prescreened as potentially eligible (i.e., number of first prenatal visits with age eligible women occurred during each day at that location).

An example patient listing table for this scenario appears below in Table 2.1. The NCS data collectors created a record for each woman who spoke to them, whether or not she completed the eligibility screener. In the example below, four women spoke to NCS staff about the Study that day. Of those women, three agreed to complete the eligibility screener, and one refused (#3 in the list). The SC obtained the PII through the eligibility screener from the women who completed the screener.

At the end of the day, the provider reported that there were a total of 30 patients who were seen by practitioners, and six of those were age-eligible, first prenatal visits. The SC created 26 “dummy” records to represent the total number of patients. All 30 records have a ‘Y’ for sampled\_person, because in Method 2, all women who come in during the sampled time period are sampled.

The first two dummy records (#5 and #6) represent the two potentially eligible women who did not stop by the table. The provider reported six first prenatal visits, but only four stopped by the table. The six records representing the age-eligible, first prenatal visit women have their eligibility criteria filled in. The other 24 dummy records do not have eligibility criteria

completed, as the provider staff did not report why these patients were ineligible.

**Table 2.1: Temporal Subsampling Only (Method 2), Scenario 1**

Person ID	Date of Visit	First Prenatal Visit? (Yes/No)	Pregnant? (Yes/No)	Age Eligible? (Yes/No)	In County? (Yes/No)	Prescreen Status	Provider Intro Outcome	Sampled Person (Yes/No)	Person First Name	Person Last Name	Phone Number
1	6/1/2012	Y	Y	Y		In	1 Female to contact SC	Y	Jane	Doe	555-555-1212
2	6/1/2012	Y	Y	Y		In	1 Female to contact SC	Y	Teresa	Garcia	555-555-1313
3	6/1/2012	Y	Y	Y		In	1 Female to contact SC	Y	REFUSED	REFUSED	REFUSED
4	6/1/2012	Y	Y	Y		In	1 Female to contact SC	Y	Mary	Smith	555-555-1515
5	6/1/2012	Y	Y	Y		In	3 (Female refused to allow SC to contact her)	Y			
6	6/1/2012	Y	Y	Y		In	3 (Female refused to allow SC to contact her)	Y			
7	6/1/2012					Out	-7 (Not Applicable)	Y			
8	6/1/2012					Out	-7 (Not Applicable)	Y			
9	6/1/2012					Out	-7 (Not Applicable)	Y			
10	6/1/2012					Out	-7 (Not Applicable)	Y			
11	6/1/2012					Out	-7 (Not Applicable)	Y			
12	6/1/2012					Out	-7 (Not Applicable)	Y			

....30	6/1/201 2					Out	-7 (Not Applicable)	Y			
--------	--------------	--	--	--	--	-----	---------------------	---	--	--	--

**Key:**

Information from provider	Implied by prescreening	SC creates information	Obtained during eligibility screener
---------------------------	-------------------------	------------------------	--------------------------------------

*Method 2 Scenarios (continued)*

2. A large public health clinic with a subsampling rate is 0.125 (1 in 8).

Westat will help the SC randomly select two weeks for recruitment. The clinic will allow the NCS recruiter to talk to their patients while waiting for appointments. The provider cannot prescreen their patients or provide a list of patients. The recruiter will need to keep a running count of all women who come in for appointments, whether they agree to talk to the NCS or not. Because prescreening will not be done, all women (pregnant or not) will need to be listed. This method may require two data collectors; one to sit with a view of the check-in desk in order to get an accurate count of the number of patients seen each day and one to interact with patients.

A sample patient listing table appears below in Table 2.2. This shows the first nine patients who came in that morning. There is a record for every woman who came in for an appointment. Records for women who do not talk to the recruiter could be created in real time, or at the end of the day using the total patient count. In the example below, records were created as the women had their appointments. Five of the first nine patients that day agreed to talk to the recruiter (#1, #2, #5, #6, and #8.) Of them, four agreed to complete the eligibility screener, and one refused (#5). Because there is no systematic sampling of women that needs to take place in method 2, it is possible to complete the eligibility screener upon first talking to each woman.

**Table 2.2: Temporal Sampling Only (Method 2), Scenario 2**

Person ID	Date of Visit	First Prenatal Visit? (Yes/No)	Pregnant? (Yes/No)	Age Eligible? (Yes/No)	In County? (Yes/No)	Prescreen Status	Provider Intro Outcome	Sampled Person (Yes/No)	Person First Name	Person Last Name	Phone Number
1	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	Y	Jane	Doe	555-555-1212
2	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	Y	Teresa	Garcia	555-555-1313
3	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	Y			
4	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	Y			
5	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	Y	REFUSED	REFUSED	REFUSED
6	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	Y	Mary	Smith	REFUSED
7	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	Y			
8	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	Y	Anna	Carpenter	555-555-1616
9	6/1/2012					No Prescreen Performed	-7 (Not Applicable)	Y			

Key:

Information from provider	Implied by prescreening	SC creates information	Obtained during eligibility screener
---------------------------	-------------------------	------------------------	--------------------------------------

*Method 2 Scenarios (continued)*

**3.** A small location with a subsampling rate of .5.

This location serves mostly private insurance patients and does not accept walk-in appointments. They only have 2-3 first prenatal visits each week. As the subsampling rate is 0.5, NCS will recruit patients every other week. The office manager at the office will call the SC on Fridays before selected weeks and give the appointment times of their new prenatal patients for the following week. The data collector will come to the office at those times. During an appointment, the doctor or nurse will introduce the NCS to the patient, and ask her if she would like to meet with the data collector. If the patient agrees, the data collector will then attempt to complete the eligibility screener with the patient, or set up an appointment to do it at a later time. At the end of the week, the location will report: how many patients they saw that week, how many of those were pregnant, and if any first natal visits were added since the beginning of the week.

A sample patient listing table appears below in Table 2.3. In this week, the location alerted the SC to three first prenatal visits. Of the three patients, one agreed to talk to the data collector (#2); one told the location she would think about it and get back to the location staff (#1) and one refused to talk to the data collector (#3).

At the end of the week, the location reports that they had a total of 38 patients, including the 3 first prenatal visits. Of those visits, 15 were not obstetrics patients. There were 20 pregnancy visits that were not first visits. The SC created 35 dummy records for the 35 ineligible patients. Records #4-23 represent the 20 pregnancy visits. The values for Pregnant are 'Y' and 'N' for First Prenatal Visit. Records #24-38 represent the 15 visits from non-pregnant patients. The values for Pregnant and First Prenatal Visit are 'N'.

**Table 2.3: Temporal Sampling Only (Method 2), Scenario 3**

Person ID	Date of Visit	First Prenatal Visit? (Yes/No)	Pregnant? (Yes/No)	Age Eligible? (Yes/No)	In County? (Yes/No)	Prescreen Status	Provider Intro Outcome	Sampled Person (Yes/No)	Person First Name	Person Last Name	Phone Number
1	6/4/2012	Y	Y	Y		In	4 = Female undecided about SC Contact	Y			
2	6/5/2012	Y	Y	Y		In	2 = Female Agreed to SC Contacting Her	Y	Anna	Carpenter	555-555-1616
3	6/7/2012	Y	Y	Y		In	3 = Female Refused to Allow SC to Contact Her	Y			
4	6/8/2012	N	Y	Y		Out	-7 (Not Applicable)	Y			
5	6/8/2012	N	Y			Out	-7 (Not Applicable)	Y			
6	6/8/2012	N	Y			Out	-7 (Not Applicable)	Y			
...	6/8/2012	N	Y			Out	-7 (Not Applicable)	Y			
23	6/8/2012	N	Y			Out	-7 (Not Applicable)	Y			
24	6/8/2012	N	N			Out	-7 (Not Applicable)	Y			
....	6/8/2012	N	N			Out	-7 (Not Applicable)	Y			
38	6/8/2012	N	N			Out	-7 (Not Applicable)	Y			

Key:

Information from provider	Implied by prescreening	SC creates information	Obtained during eligibility screener
---------------------------	-------------------------	------------------------	--------------------------------------



### **Method 3: Temporal and Systematic Sampling**

This combines method 1(systematic subsampling) and method 2 (temporal subsampling). Entire weeks would be subsampled (1 in X weeks). Like method 2, the value of X is chosen so that at least two weeks are included. In addition to sampling temporally, women will be subsampled from the lists built in the selected weeks.

This method may be more likely to be under consideration in provider locations with low sampling rates (i.e. in larger locations).

### **Method 4: Two Levels of Temporal Sampling**

This is similar to method 2 (temporal subsampling), in that a 1 in every X is selected. However, for provider locations with a lower subsampling rate, we may need to sample less than entire weeks. Within selected weeks, 1-4 days are subsampled. The days of the week should be different from week to week in order to reduce the potential for selection bias due to differential patient flow throughout the week. For example, if 3 weeks are selected, and 3 days a week are needed, the following scenario is an example of one possible outcome:

- 1st subsampled week, the list might be built Monday-Wednesday,
- 2nd week, Tuesday-Thursday, and
- 3rd week Wednesday-Friday.

### **Using provider logistics to determine method of patient sampling**

These scenarios are intended to arm SCs with the knowledge necessary to discuss a preferred sampling and recruitment strategy with each selected provider location. Before discussing the method for selecting patients within a location, you should have an ideal method already in mind for that provider as well as a backup method if the provider does not agree to the first one for some reason. It is not advisable to simply give a list of all of the options to a location; it will overwhelm them and may jeopardize their participation.

## **Patient sampling rate**

For small and medium sized locations, those with sample rates of .125 or more, any of the methods of selecting women within the location may be employed. Because methods 1 and 2 are not as complex as 3 and 4, they may be easier for the provider location staff to understand and make them more agreeable to participating.

Larger locations will tend to have small sampling rates of women due to the nature of the sample design for PBS. Methods 3 and 4 are more conducive to these larger locations with lower sampling rates for a couple of reasons. First, while building a complete list of patients during the field period might be feasible (method 1), it is not practical because you will be selecting such a small ratio of women. Method 2 is not possible to employ when the sampling rate is below .125. Methods 3 and 4 allow you to minimize the amount of time in the location and the amount of work the location staff and SC staff need to do. *Appendix 1: Methods for Sampling Women within Selected Providers* contains examples of selection rates. These tables further illustrate how important the patient sampling rate is when determining the sampling method within a location.

## **Eligibility screening**

In most locations, administering the eligibility screener to selected women will take place retrospectively, at a separate time after the first prenatal visit. However, in specific instances of methods 2 and 4, real-time eligibility screening may take place. When the temporal sampling of women brings the sample rate to 1, in other words taking all of the women who come in during the selected time period(s), real-time screening is a viable option. This may introduce further efficiencies and allow SCs to get higher response rates by the added legitimacy of conducting the eligibility screener in the provider office.

## **Patient List**

The creation of the list of patients from each of the sampled locations brings up several operational considerations. First, who is actually compiling the list? This may be produced by Provider staff at the location or could potentially be created by NCS staff working at the location. If the list is being built by NCS staff, some form of temporal sampling would be preferable in order to save on resources. Second, the location may not be able to create a list easily due to the lack of electronic record keeping. In these cases, it may be best for NCS staff to generate the lists themselves. If the Provider staff at the location is willing and able to generate the list of women, any of the methods discussed in this memo will work.

It is important to keep in mind that the patient list is highly dependent on two other factors: prescreening and release of PII. It is likely that several iterations of patient lists may be passed back and forth between the provider location and NCS staff.

Preliminary iterations may include dummy IDs for each of the women, and only after obtaining permission to release PII will a list be created with actual names on it.

### **Prescreening**

All methods will be more efficient if prescreening can be done by the location or SC staff. The amount of prescreening does not affect the sampling rate at all, but it does reduce the number of women who need to be approached to complete the eligibility screener. For example, take two lists of 100 women, one that had not been prescreened and one that had been prescreened. For the first list (not prescreened), after applying the sampling rate of women (0.5 in this example), there will be 50 selected women who will need to complete the eligibility screener. Of these 50, 8 might meet all four eligibility criteria. The second list was prescreened for age, pregnancy, and first prenatal visit at that location, resulting in a final list of the 20 women. Applying the same sampling rate (.5) will result in list of 10 women who need to complete the eligibility screener, and 8 of which will be determined to be eligible. The prescreening reduced the amount of effort required to screen women for eligibility, but did not affect the final result.

### **Release of PII**

In most locations, location staff will need to ask women permission to release their PII to the SC. Regardless of the method of selecting women in the location, it will be best if permission to release PII is gained at the time of the first prenatal visit. If location staff do not want to ask permission from women who may not be sampled into the study, methods 2 and 4 would be preferable, as the staff would know every patient during the selected time period would be in sample.

Questions/Comments: If you have any questions or concerns about this memo or are in need of Field Support assistance, please contact the NCS Help Desk at [support@nationalchildrensstudy.gov](mailto:support@nationalchildrensstudy.gov) or call toll free at 866-817-7033.