Appendix C
Sampling

C.1 Definition

The population of interest for the Puerto Rico Pregnancy Risk Assessment Monitoring System - Zika Postpartum Emergency Response (PRAMS-ZPER) is all women who are residents of Puerto Rico and who delivered a live-born infant in Puerto Rico during the surveillance period, and their available male partners. To draw a sample from this population, a sampling frame must be identified. Since 99.17% of the births in Puerto Rico occur in the hospital setting, PRAMS-ZPER will employ hospital-based data collection. The operational target population is all resident women delivering a live birth in selected field sites (Puerto Rican hospitals) during the surveillance period, and their available male partners.

In practice, the hospital-based sampling frame is identified by the hospital's delivery log. In most hospitals the delivery log represents a complete and accurate account of all births occurring in the hospital. It will be important to insure that the delivery log in each participating hospital is reliable and kept up-to-date. If this is not the case, alternate procedures to identify women who are eligible for PRAMS-ZPER will be developed specifically for that field site.

PRAMS-ZPER data collection will take place in all hospitals in Puerto Rico with 100 or more births per year. According to 2016 birth data obtained from Demographic Registry, 34 hospitals meet this eligibility criteria. However, one qualifying facility in the Metro region will not be included as a PRAMS-ZPER field site as their delivery ward closed down during September 2016. Due to the effects of hurricane María, two additional hospitals (one in Caguas and another in Aguadilla) will not be included since the maternity wards will not be working throughout the sampling period due to the damages suffered from the hurricane. It is assumed the births that would have occurred in these hospitals will take place at other hospitals participating in PRAMS-ZPER. Lastly, one hospital in the metro area did not agree to participate in the survey, allotting a total of 30 hospitals to PRAMS-ZPER 2.0 which will represent 96.2% of all births in Puerto Rico. Data collection will take place over a 2-month period in the fall of 2017. The overall sample size for ZPER 2.0 is 1470 women and 1,000 male partners.

Subsequently, all participating women will be enlisted to complete the PRAMS-ZPER Telephone Follow-up Survey.

C.2 Adjustments to the Sampling Frame

C.2.a Inclusions and Exclusions. Exclusions to the PRAMS-ZPER sampling frame may be implicit or explicit. Because of the definition of the PRAMS-ZPER targeted population and the use of the hospital delivery logs as the sampling frames, certain mothers will *implicitly* be excluded from eligibility in the ZPER sample. An implicit exclusion is any restriction inferred by the definition of the targeted population (i.e., non-residents of Puerto Rico and women who deliver at home or at hospitals with fewer than 100 births per year) or the choice of the hospital delivery log as the PRAMS-ZPER sampling frame (i.e., stillbirths and fetal deaths). All other exclusions arise from concerns or operational difficulties in sampling certain types of births, and are termed *explicit* exclusions.

i. ***Stillbirths and Fetal Deaths.*** By definition, the targeted population of PRAMS-ZPER is limited to pregnancies resulting in a live-born infant.
ii. ***Multiple Gestation Pregnancies.*** Mothers with a multiple gestation regardless of the order will be included in the sample if at least one infant is delivered, alive, in the sampling window. Since the survey questions are about the mother and her pregnancy and not specifically about the baby, the mother will complete one survey for this pregnancy, regardless of how many babies were delivered.
iii. ***Mothers Discharged Early From the Hospital or Otherwise Missed in the Hospital.*** For hospital-based supplementation, mothers discharged early from the hospital or otherwise missed in the hospital are not to be excluded from the sample. Nonetheless, no additional procedures for locating and contacting them will be developed. On the other hand, mothers transferred to another hospital participating in the PRAMS-ZPER survey will not be excluded. Location and follow-up of transferred mothers will be arranged.

iv. ***Deceased Infants.*** Mothers of deceased infants will be excluded from the sample. The Neonatal Mortality Rate for 2016 births in Puerto Rico is 5.33. Based on 2016 neonatal mortality rates in Puerto Rico, approximately 16 infant deaths will occur during the study period. This number is too small to have any impact on the sample size calculations.

v. ***Mothers with Pregnancy Complications.*** Mothers who are ill or suffering from complications of pregnancy and delivery will not be excluded. Contact with these women may need to be delayed until their condition has improved. However, they should be approached to complete the survey before they are discharged from the hospital.

vi. ***Births at San Juan Municipal Hospital –*** Since this hospital did not agree to participate, births from this hospital are excluded from the study. With approximately 100 births per month, an adjustment was made to the sample size to compensate for the births that will be missed at this hospital.

C.3 Sampling Plan

C.3a Sampling Scheme for PRAMS-ZPER. For PRAMS-ZPER surveillance, an island-wide proportional sample will be drawn.

C.3b Determining Overall Sample Size. Required sample sizes for the questionnaire are determined in relation to the given proportion that is being estimated, at a given level of precision, and with a given level of statistical confidence. For specified levels of precision and confidence, the sample size required is at its maximum when the advance estimate (the number used in sample size calculations) of the proportion being estimated equals 0.50. PRAMS-ZPER data are used in estimates of proportions of risk factors that range from common (such as delivery paid for by Medicaid) to rare (such as a confirmed Zika diagnosis). Using 0.50 in sample size calculations leads to sufficiently large sample sizes, whatever the true population proportions are for the various risk factors.

For the PRAMS-ZPER sampling plan, a sample size of about 1068 (n = 1068) is necessary to estimate a prevalence for a dichotomous variable with a reasonable precision of 3% and a confidence level of 95%, assuming an infinitely large population size (N). When the assumption of an infinitely large population will be violated, it is appropriate to apply the finite population correction (FPC). The FPC will reduce the desired sample sizes in such cases without compromising the precision of the estimates.

The formula for FPC is:

$$\boldsymbol{FPC} = \frac{\boldsymbol{n}}{(\mathbf{1} + (\boldsymbol{n}/\boldsymbol{N})}$$

Where, $\boldsymbol{n = Desired\ sample\ size}$

$\boldsymbol{N = Population\ Size}$

Not all sampled mothers will participate in the study. Thus, actual sample sizes must be larger than theoretically needed to achieve a given level of statistical power. Based on the estimated response rates, the sample size will be inflated to ensure an adequate number of responses for analysis. Based on previous ZPER hospital-based surveillance in Puerto Rico, an 81% response rate is assumed.

Births in Puerto Rico have been steadily declining in recent years. The most recent birth data by hospital available is for 2016. Since births have continued to decline, a sampling rate based on 2016 birth distributions would not achieve the desired sample size. Therefore, it is necessary to account for the declining birth rate. Table C.1 describes the drop in birth rates from 2007 to 2016. Based on actual birth data for the first 5 months of 2016 and 2017 in hospitals with 100 or more births per year, an adjustment factor of 1.1362 will be used to estimate the number of 2017 births. Due to one hospital not participating, the expected number of births was further reduced to exclude births occurring at that hospital.

Table C.1 Annual percent change (APC) for declining births in Puerto Rico from 2007 to 2016.		
Year	**Annual Births**	**% Decline (APC)**
2007	46,750	
2008	45,689	-2.27%
2009	44,830	-1.88%
2010	42,203	-5.86%
2011	41,133	-2.54%
2012	38,974	-5.25%
2013	36,580	-6.14%
2014	34,493	-5.71%
2015	31,233	-9.45%
2016	28,321	-9.32%
Jan – May 2016	*21,158**	
Jan – May 2017	*18,276**	*-13.62%*
Note. Values in *italics* are estimates. Data for 2016 is preliminary and was obtained from the Puerto Rico Demographic Registry in 06/2017. *Births in hospitals with 100 or more births per yeat.		

C.3*c* Steps for Establishing the Sample Rates

1. **Establish the distribution of births in Puerto Rico by hospital.** Obtain a list of births by hospital. This list was provided by the Puerto Rico Health Department for 2016 births. Determine which hospitals have a sufficient number of births to support PRAMS-ZPER surveillance.

2. **Select the hospitals where data collection will occur.** Criteria for hospital selection should be defined. For ZPER, all hospitals with at least 100 births per year will be included. A total of 34 hospitals meet this eligibility criteria for PRAMS-ZPER; nonetheless, one qualifying facility in the Metro region was as their delivery ward closed down during September 2016. Due to the effects of hurricane María, two additional hospitals (one in Caguas and another in Aguadilla) will not be included since the maternity wards will not be working throughout the sampling period due to the damages suffered from the hurricane. It is assumed the births that would have occurred in these hospitals will take place at other hospitals participating in PRAMS-ZPER. Lastly, one hospital in the metro area did not agree to participate in the survey. The adjustment leaves a total of 30 hospitals for PRAMS-ZPER 2.0 which will represent 96.2% of all births in Puerto Rico. Table C.4 shows a list of the 30 selected field sites for PRAMS-ZPER 2.0 and the total number of births that occurred in that facility during 2016.

3. **Calculate the number of eligible mothers.** Mothers giving birth to double or higher plurality result in multiple births, but only one eligible mother. The multiple birth rate is 2.02% of total births. The number of distinct mothers for Puerto Rico during 2016 was determined (28,027) for each region. Thus the number of eligible mothers can be estimated as 98.96% of the total births.

4. **Adjust for estimated declines in the birth rates from 2016 to 2017**. The adjustment factor is calculated by dividing the total number of births during 2016 by the estimated births for 2017. We calculated the adjustment factor from the data in Table C.1. The adjustment factor is 1.1362 for 2017 births. Divide by 1.1362 to get adjusted eligible mothers for PRAMS-ZPER 2.0.

5. **Determine the desired number of respondents in the sample.** This number will be based in part upon costs and resources. It was chosen to be 1068, as an estimate of a proportion based upon 1068 respondents will have a 95% confidence interval of +/- 3%.

6. **Compute the Finite Population Correction** (FPC), if applicable. See formula in section *C.3b*.

7. **Estimate the completion rate of hospital-based data collection.** Based on PRAMS-ZPER previous hospital-based survey implementation using the same in-hospital methodology, a total of 2,364 women completed the survey from 2,933 eligible births (80.60%). An 81% response rate is assumed across all strata. Divide the FPC Corrected Sample Size by the estimated response rate to determine the estimated sample size adjusted for non-respondents (final sample size).

8. Complete Table C.2 using the result of steps 3 through 7 to fill in the appropriate columns.

   Carry the adjusted population size and estimated adjusted sample size from Table C.2 to Table C.3.

9. **Compute the population size for the 6-week surveillance period**. Determine the annual birth distribution by weeks, and establish the percent of births observed per week. Later determine the percent of births for the surveillance period to validate that there is no variation in births during the proposed sampling period. The annual births will be divided into weeks, thereby in order to calculate the population size for the 6-week surveillance period the annual population size must be divided by 52 (number of weeks in a year) and then multiply by 6.

10. **Establish the sampling fraction.** The sampling fraction is the proportion of the sample that will be included in the sample. The sampling fraction is calculate by dividing the desired sample size (*ZPER Estimated Adjusted Sample Size*) by the population size (*ZPER Estimated Population Size during the 3 Month Period*). See formula for sampling fraction below:

$$\boldsymbol{Sampling\ Fraction\ (f) = \frac{n}{N}}$$

Where, $\boldsymbol{n = Desired\ sample\ size}$
$\boldsymbol{N = Population\ Size}$

11. **Calculate the Operational Sample Size.** The operational sample size is the expected sample size based on the number of births actually occurring. The operational sample size is calculated by multiplying the Estimated Sample Size (*ZPER Estimated Adjusted Sample Size*) times the sampling fraction (*f*).

12. **Determine the sampling fraction in days.** The 6-week sampling period (November 1, 2017 to December 12, 2017) has a total of 42 days. From the sampling fraction, determine the number of days over the 6 week (42 day) surveillance period during which sampling will be conducted.

13. Using the number of days during which sampling will be conducted, determine the intervals for sampling. Using the example of 2 out of every 15 days, the two days should be randomly chosen at each hospital. For nearby hospitals, the sampling days can be shifted by one or two days to distribute the workload. Care should be taken to vary the days of the week that sampling occurs throughout the surveillance period. Specific sampling schedules for each field site will be provided by CDC.

14. Note: After the study was underway, it appeared the number of actual births was lower than expected and the study period was extended by one week. Thus the final study period was 49 days from (November 1, 2017 to December 19, 2017) and the final sample size increased from 1260 to 1540. The same sampling rate f=.475 was applied in the extra week.

	Total Number of Live Births	Live Births in Eligible Hospitals	Total ZPER Eligible Population of Mothers (adjusting for non-residents and Multiple Births) in Eligible Hospitals	ZPER Population Size Adjusted for Projected 2017 decline in births	Estimated Unadjusted Sample Size	FPC Corrected Sample Size	Estimated 6-week Sample Size for Non-Respondents (81% based on ZPER 1.0)
TOTAL	**28,321**	**27,680**	**26,100**	**24,046**	**1068**	**1021**	**1260**

Table C.3 PRAMS-ZPER 2.0 Sampling Fractions and Estimated Sample Sizes (extended to 7 weeks)						
	ZPER Population Size Adjusted for Projected 2017 decline in births	**ZPER Estimated Population size During 7 week Study Period**		$f = \frac{n}{N}$	**Operational Sample Size**	***f* in days (number of days to sample out of the 49 days)**
Total	**24,046**	**3,237**		**.475**	**1,540**	**20**
Source: Demographic Registry 2016 preliminary data.						

Table C.4 PRAMS-ZPER 2.0 Selected Field Sites, Hospitals		
Region	**Field Site Name (Hospital Name)**	**Total Births During 2016**
AGUADILLA	HOSPITAL SAN CARLOS BORROMEO	554
ARECIBO	DOCTORS CENTER MANATI	1,019
ARECIBO	HOSP. PAVIA ARECIBO/ CAYETANO COLL Y TOSTE	556
ARECIBO	MANATI MEDICAL CENTER (DR. OTERO LOPEZ)	2,024
BAYAMON	DOCTORS CENTER BAYAMON	302
BAYAMON	HOSPITAL HERMANOS MELENDEZ	1,213
BAYAMON	HOSPITAL HIMA SAN PABLO BAYAMON	1,218
CAGUAS	HOSPITAL HIMA SAN PABLO CAGUAS	1,691
CAGUAS	HOSPITAL MENONITA AIBONITO	511
CAGUAS	HOSPITAL MENONITA CAGUAS	405
CAGUAS	HOSPITAL MENONITA CAYEY	1,107
FAJARDO	CARIBBEAN MEDICAL CENTER	343
FAJARDO	HOSPITAL HIMA SAN PABLO FAJARDO	259
MAYAGUEZ	HOSPITAL BELLA VISTA MAYAGUEZ	460
MAYAGUEZ	HOSPITAL DE LA CONCEPCION	550
MAYAGUEZ	HOSPITAL METROPOLITANO SAN GERMAN	274
MAYAGUEZ	HOSPITAL PEREA MAYAGUEZ	583
MAYAGUEZ	HOSPITAL SAN ANTONIO (MUNICIPAL MAYAGUEZ)	955
METRO	ASHFORD PRESBYTERIAN HOSPITAL	2,460

METRO	HOSPITAL AUXILIO MUTUO	1,018
METRO	HOSPITAL PAVIA HATO REY	980
METRO	HOSPITAL PAVIA SANTURCE	1,228
METRO	HOSPITAL UNIVERSITARIO	1,108
METRO	HOSPITAL UPR CAROLINA FEDERICO TRILLA	600
PONCE	HOSPITAL DAMAS	961
PONCE	HOSPITAL DR. PILA	371
PONCE	HOSPITAL EPISCOPAL SAN LUCAS GUAYAMA	457
PONCE	HOSPITAL METROPOLITANO DR. TITO MATTEI	205
PONCE	HOSPITAL SAN CRISTOBAL	414
PONCE	HOSPITAL EPISCOPAL SAN LUCAS I Y II	1,810
	TOTAL BIRTHS IN ZPER SELECTED FACILITIES	**27,680**
Source: Demographic Registry 2016 preliminary data.		

C.4 Selection of Sample

Proportional sampling is used for drawing the sampling schedule based on the time of birth. The time and date of birth is written on the hospital delivery log. All births that fall within the pre-established sampling time intervals are selected for the study provided they do not satisfy an exclusion criterion. Sampling intervals will consist of complete days (midnight to midnight) for ease of selection. The sampling schedule is designed to be balanced by weeks in the surveillance period and days within each week.

Based on the desired sample size and the number of live births occurring at each hospital, sampling fractions can be computed. The length of the sampling interval is determined from the sampling fractions. For a multiple birth, the mother is selected only once. The selection procedures must satisfy the probability requirements of the sample. The sample is chosen so that each record has an equal probability of being selected. Based on these probabilities, weights can be determined for island-wide estimates.

Because of shorter hospital delivery stays and earlier discharges of mothers, the hospital delivery log must be frequently monitored during defined sampling intervals. No more than 24 hours should lapse between the beginning of the sampling interval and when the delivery log is checked. Similarly no more than 24 hours should lapse between checks of the delivery log within a sampling interval.

C.5 Paternal In-hospital Sampling

The target sample for the paternal survey is all husband/partners of mothers selected for the maternal sample. Since the maternal sample is representative of all Puerto Rico births, the paternal sample will also be representative of all Puerto Rico births. Fathers will be selected based on their presence in the hospital during the time the surveyor is contacting mothers or subsequent time prior to the mother being discharged. Only fathers of qualifying mothers will be approached and

attempts will be made to contact all fathers of sampled births. Additional procedures can be developed to address fathers of qualifying infants that are not present during the interviewer's initial visit.

The sample size for paternal surveying is expected to be 60-80% of the sample size determined for mothers. The sample size determined for PRAMS-ZPER 2.0 is 1540, meaning father sampling will be between 924 and 1,232. Table C.5 shows the sample size estinates for the paternal survey.

Sample size for fathers will be dependent on encountering the father in the hospital during the period in which the mother is eligible to be interviewed. Completion of paternal survey does not replace the need to promote postpartum women's participation. In the event that women decide to be non-participants, available fathers may still be considered for participation. Paternal in-hospital data collection modes will include paper/pen, tablet, or laptop.

Table C.5
ZPER Estimated Sample Sizes for Paternal Survey

	PRAMS-ZPER Operational Sample Size (from Table C.3)	**Paternal Estimated Sample Size (60%)**	**Paternal Estimated Sample Size (80%)**			
Total	**1540**	**924**	**1232**			

C.6 Educational Intervention Sampling

The PRAMS-ZPER Educational Intervention (EI) will be provided to all eligible women and available partners, regardless of their participation.

C.7 Telephone Follow-Up Survey Sampling

C.7*a* **Definition.** The population of interest for the Puerto Rico PRAMS Zika Postpartum Emergency Response (PRAMS-ZPER) 2.0 Telephone Follow-up survey is Puerto Rican resident women who delivered a live-born infant in Puerto Rico and participated in the PRAMS-ZPER 2.0 hospital-based survey in the fall of 2017.

The PRAMS-ZPER 2.0 Telephone Follow-up Survey (TS) will target all in-hospital study maternal participants. The Puerto Rico Department of Health will match information from PRAMS-ZPER respondents with the birth certificate record of their baby. The list of original study participants who can be matched to a birth certificate record will serve as the sample. The sample will include women who gave birth during the 7-week study period from November 1 thru December 19, 2017. Since the ZPER follow-up survey will be focused on postpartum behaviors, recall should not be a concern. Thus, there is no need to impose age limits on the infants included in the study.

C.7*b* **Inclusions and Exclusions.** Exclusions to the PRAMS-ZPER 2.0 Telephone Follow-up Survey target sample may be implicit or explicit. Because of the definition of the PRAMS-ZPER targeted population and the need to match PRAMS-ZPER data with birth certificate data to identify birth to PRAMS-ZPER

participants, certain mothers will implicitly be excluded from eligibility in the PRAMS-ZPER follow-up sample. An implicit exclusion is any restriction inferred by the definition of the targeted population (i.e., women who did not participate in PRAMS-ZPER or PRAMS-ZPER respondents who gave birth outside the stated time period) or the choice of the list of PRAMS-ZPER respondents matched to a birth certificate record as the PRAMS-ZPER follow-up sample (i.e., records that could not be matched to a birth certificate). All other exclusions arise from concerns or operational difficulties in sampling certain types of births, and are termed explicit exclusions.

i. ***Deceased infants.*** Mothers of infants who have died will be included in the sample if they completed the PRAMS-ZPER 2.0 hospital survey.
ii. ***Multiple Gestation Pregnancies.*** Mothers with a multiple gestation regardless of the order will be included in the sample. In PRAMS-ZPER, the mother was asked to respond to the infant questions for just one of her randomly selected infants. The same randomly selected infant will be eligible for the PRAMS-ZPER follow-up survey. All other infants from the pregnancy will be excluded.
iii. ***Mothers with Zika-affected or special needs infants*.** Mothers whose infants have microcephaly or other birth defects or special health conditions will not be excluded. If they are contacted, the interviewer should be aware of extra sensitivity of questions related to the baby's health.

C.7*c* **Sampling Plan.** The PRAMS-ZPER 2.0 In-Hospital sampling design was an island-wide proportionasl sample. Similarly, the PRAMS-ZPER 2.0 Telephone Follow-up Survey will utilize a proportional sample.

The sampling plan is designed so that island-wide inferences about prevalence rates for maternal behaviors and knowledge of Zika can be estimated with sufficient precision.

C.7*d* **Determining Overall Sample Size.** No sampling is needed for the PRAMS-ZPER TS. All participants in the ZPER hospital study that are matched with a birth certificate record are included. PRAMS-ZPER participants must be matched to their infant's birth certificate in order to collect current contact information. The Puerto Rico Department of Health is performing the linkage based on 5 fields, mother's name, mother's date of birth, baby's date of birth, hospital of delivery, and method of delivery. Based on PRAMS-ZPER 1.0 activities, we expect a 99% matching rate. The PRAMS-ZPER participants who are successfully matched to their infant's birth certificate record form the sample for the PRAMS-ZPER 2.0 Telephone Follow-Up survey.

Based on the 77% response rate achieved for the ZPER 1.0 telephone survey, it is expected that there will be approximately 1138 complete telephone surveys.

Table C.6 ZPER 2.0 Telephone Survey Expected Sample Size

	PRAMS-ZPER Participants n	**PRAMS-ZPER Participants Adjusted for Matching** n	**Expected number of completed telephone surveys (based on 77% response rate)**
Total	1493	1478	1138