Describing the PRAMS Sample Design for SUDAAN, SAS Complex Survey, SPSS Complex Samples Modules, STATA, and R

CDC PRAMS has performed the following actions while constructing an analysis dataset for external researchers. External researchers may skip this step and proceed directly to SUDAAN/SAS/SPSS/STATA/R analysis (step II below).

I. Combine the single-year/single-state weighted PRAMS SAS datasets into a single analysis dataset

A. Restrict the dataset to respondents only:
IF INQX=1;

B. Combine the state stratification scheme (variable STRATUMC) and sample year (variable NEST_YR) into a single variable:
SUD_NEST = (STRATUMC*10000) + NEST_YR;

C. Sort the dataset by new variable SUD_NEST:
PROC SORT DATA=<pramsdata>; BY SUD_NEST;

II. Describe the PRAMS Sample Design to the analysis software

A. SUDAAN

1. Use the following statements when using SUDAAN modules:

```
proc crosstab data=<pramsdata> design=strwor;
nest sud_nest;
totcnt totcnt;
samcnt samcnt;
weight wtanal;
```

B. SAS Complex Survey

1. Construct a new SAS dataset describing the population count for each sampling stratum. The new dataset will have one observation per sampling stratum, and two variables: SUD_NEST (the sampling stratum) and _TOTAL_ (the population count).

   One method for creating this dataset is as follows:

```
Data totals_for_sas;
Set <pramsdata> (keep= sud_nest totcnt);
By sud_nest;
If first.sud_nest;
_total_=totcnt;
Keep sud_nest _total_;
Run;
```

2. Use the following statements when using SAS Complex Survey modules:

```
proc surveyfreq data=<pramsdata> nomcar total=totals_for_sas;
strata sud_nest;
weight wtanal;
```

C. SPSS Complex Samples

1. To describe the PRAMS sample design, open your PRAMS dataset, and choose Analyze/Complex samples/Prepare for Analysis/Create a plan file

2. From the 'Stage 1: Design Variables' screen:
   a. Select variable SUD_NEST and move it into the 'Strata' box
   b. Leave the 'Clusters' box blank
   c. Select variable WTANAL and move it into the 'Sample Weight' box

3. From the 'Stage 1: Estimation Method' screen:
   a. Choose 'Equal WOR'

4. From the 'Stage 1: Size' screen:
   a. For 'Units', choose 'Population Sizes'
   b. Choose the 'Read values from variable' box, select variable TOTCNT, and move it into the box

5. From the 'Stage 1: Plan Summary' screen:
   a. Choose 'No, do not add another stage now'

6. Save the plan file, and proceed with appropriate analysis

D. STATA

Use the following statement when using STATA:

svyset id [pweight=wtanal], strata(sud_nest) fpc(totcnt)

Or use

svyset *_n [pweight=wtanal], strata(sud*_nest) fpc(totcnt)

E. R software package

When using R, load the "survey" package and create a survey design object with the following function call:

prams.svy <- svydesign(ids = ~0, strata = ~SUD_NEST, fpc = ~TOTCNT, weights = ~WTANAL, data = <pramsdata>)

or, equivalently

prams.svy <- svydesign(ids = ~0, strata = ~STRATUMC + NEST_YR, fpc = ~TOTCNT, weights = ~WTANAL, data = <pramsdata>)

The R survey package provides a function "svyciprop" to compute asymmetric confidence intervals using the logit method.