Washington State
Pregnancy Risk Assessment Monitoring
System (PRAMS) Data Analysis Guide

To request this document in another format, call 1-800-525-0127. Deaf or hard of hearing customers, please call 711 (Washington Relay) or email doh.information@doh.wa.gov.

Publication Number: 141-099

May 2024

For more information email: WAPRAMS@doh.wa.gov

Table of Contents

PRAMS Overview

Project Purpose

The overall purpose of the Pregnancy Risk Assessment Monitoring System (PRAMS) is to find out why some infants are born healthy, and others are not. Survey answers give us information about what behaviors and factors may affect the health of the mother and their new baby. This information is used to reduce infant morbidity and mortality and improve the health of mothers in Washington State (WA). It is the only surveillance system that provides data about pregnancy and the first few months after birth.

Survey History

PRAMS is an ongoing, population-based surveillance system designed to identify and monitor selected maternal experiences and behaviors that occur before and during pregnancy and during the child's early infancy among a stratified sample of women delivering a live birth. PRAMS was initiated in 1987 as part of the Centers for Disease Control and Prevention (CDC) initiative to reduce infant mortality and low birthweight.

WA joined PRAMS in 1991 and WA Department of Health (DOH) began data collection in 1993. The questionnaire is revised periodically. With each revision or new phase, some of the questions change. Although most indicators can be compared across phases, it is often easier to analyze data within a single phase. To add questions between phases, supplements can be appended to the core survey to address emerging topics of interest. For more information on questions asked during each phase, please visit: https://doh.wa.gov/data-statistical-reports/health-behaviors/pregnancy-risk-assessment-monitoring-system/researchers

Table 1: PRAMS Phases and Birth Cohort Years

Phase	Birth Cohort Years
Phase 2	1993-1995
Phase 3	1996-1999
Phase 4	2000-2003
Phase 5	2004-2008
Phase 6	2009-2011
Phase 7	2012-2015
Phase 8	2016-2022
Phase 9	2023-Present

PRAMS Population of Interest

The population of interest for PRAMS is all birthing persons who are residents of and delivered a live-born infant within WA each year. WA's vital records birth certificate file serves as the best available source of the sampling frame representing live births, including infants who have died during infancy. Infant deaths are identified before initial contact with the mother (see Figure 1 below). A separate letter is used to express sympathy for their loss. The following groups are excluded from the sampling frame:

- Out-of-state births to residents
- In-state births to nonresidents

- Missing information on the birth certificate
- Delayed or early processing of birth certificates
- Duplicate birth(s) from multiple gestation infants (only one infant is randomly selected for inclusion in sampling frame)
- Adopted infants
- Surrogate births

Survey Sampling

WA randomly samples about 2,000 live births each year. PRAMS uses a sampling approach that oversamples some subpopulations within the population. With this approach, we are assured of having an adequate sample size for analyses of subpopulations of particular interest. WA sampling approach involves a stratified random sample based on the following five race and ethnicity sub-groups: White non-Hispanic, African American non-Hispanic, Asian/Pacific Islanders non-Hispanic, Native American non-Hispanic, and Hispanic.

Since we sample at different rates by sub-population, we must apply analytic weights to our analyses to more accurately represent the WA birthing population. See *Getting to Know PRAMS Data → Analysis Weights* section below for more details.

Survey Administration

PRAMS data are collected on an annual cycle. A monthly sample is drawn among new mothers 2-6 months after they deliver their babies. The questionnaire is self-administered by mail with telephone follow-up for non-respondents. The mail and telephone surveys are available in English and Spanish. A check to identify infant deaths is completed one day before the first mailing and first phone call.

FIGURE 1:

WA PRAMS Data Collection Methodology
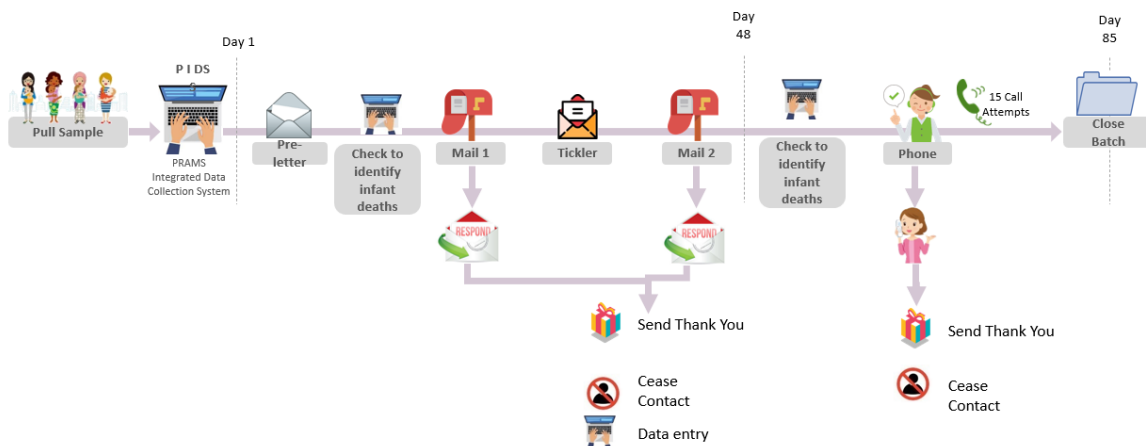
Survey Participation Rates

WA surveys approximately 2,000 women per year. CDC sets a minimum weighted response rate recommended for analyses, shown in Table 2 below. WA response rates vary by year, and we have seen a decline in response rates in recent years. Even with these declines, WA still maintains weighted response rates higher than CDC's threshold. It is important to note that these represent state-level response rates. Some sub-populations may have lower response rates, making estimates less precise. Sub-population response rates should be determined if possible, and caution should be used in interpreting results among groups with response rates lower than CDC's threshold.

Table 2: Number of women sampled and response rates for Phase 8, 2016-2022

Year*	Number Sampled	Number Respondents	Unweighted Response Rate	Weighted Response Rate	CDC Response Rate Minimum
2016	2,444	1,277	52.2%	58.0%	55%
2017	2,209	1,238	56.0%	61.4%	55%
2018	2,085	1,145	54.9%	61.5%	50%
2019	2,030	1,185	58.4%	64.1%	50%
2020	2,068	1,239	55.9%	64.0%	50%
2021	2,068	1,147	53.9%	61.0%	50%
2022	2,142	1,113	52.0%	55.1%	50%

*Response rates for previous years available upon request

Getting to Know PRAMS Data

Data Sources

PRAMS data consist of information collected from the following three data sources:

- **Birth Certificate Data.** Birth certificates serve as a source of demographic and clinical information about the sampled mother and the infant. They also provide the sampling frame and are used to weight the survey data to represent the population.
- **Operational Data.** Operational data are used to calculate response rates and to monitor the quality of operations. They are also used for analysis of PRAMS survey methodology.
- **Questionnaire Data.** Self-reported data are collected by mail and telephone. The PRAMS questionnaire serves as the principal source of maternal behavioral information for the time before, during, and after the mother's most recent pregnancy.

Survey Topics and Uses

PRAMS provides representative data not available from other sources about pregnancy and the first few months after birth. The current Phase 8 questionnaire includes questions about the following:

- Attitudes and feelings about the most recent pregnancy
- Preconception care
- Content of prenatal care
- Breastfeeding
- Cigarette, marijuana, and alcohol use
- Health insurance coverage
- Physical abuse
- Infant health and care
- Contraceptive use
- Postpartum depression
- Oral health

PRAMS data are used by state, territorial, and local agencies to plan and assess programs and policies aimed at reducing health problems among mothers and infants, as well as by researchers to investigate emerging issues in the field of maternal and child health.

Analysis Weights

As result of the PRAMS stratified sampling design, the PRAMS sample differs from the "true" distribution of these characteristics in the birthing population. To more accurately represent the population, analysis weights are calculated and applied to account for bias. Analysis of the data without using an analysis weight produces incorrect results for descriptive measures, variance estimates, and statistical tests. Components to PRAMS analysis weight include the following:

- *Selection probability*: Accounts for PRAMS stratified sampling design, where we sample at different rates by strata.
- *Unit nonresponse adjustment*: Fixes over or under representation of a particular stratum respondents. It is assumed that nonrespondents would have provided similar answers, on average, to respondents' answers for that stratum and adjustment category.
- *Nonconvergence/omissions*: Accounts for problems that occur during frame construction, such as omitted records due to late processing for birth certificates.

For more details on weighting methodology, please visit: cdc.gov/prams/methodology.htm . Application of analysis weights during analyses is described below.

Publishing Requirements

There are small number publishing requirements to ensure participant confidentiality and statistical reliability. Please note PRAMS requirements are different than WA DOH standards. Current requirements are as follows:

- Do not publish results if the number of participants is less than 30.
- Include a flag next to the result if RSE is between 25% and 30%.
- Do not publish results if RSE is greater than 30%.

Some domain analyses or analyses of rare behaviors or health outcomes may require aggregating data values, such as combining two or more years of data, age groups, or diagnoses, to achieve the minimum sample size.

Assessing and reporting the relative standard errors (RSEs) 95% and Confidence Intervals (CIs) allow the audience to understand the stability and reliability of the results and decrease the likelihood of misinterpretation due to small numbers.

Relative Standard Error (RSE)

RSEs provides a measure of reliability for statistical estimates. When the RSE is large, the estimate is imprecise, and we term such rates "unstable" or "not reliable." Large RSE merit annotation that the estimate is not reliable or refrain from publishing. PRAMS estimates should not be published if the RSE is greater than 30%. Additionally, estimates with RSEs between 25% and 30% should be published with a flag indicating potential unreliability.

You can find more details, including how to calculate RSE in WA Department of Health Agency Standards for Reporting Data with Small Numbers, pages 13-15: [https://doh.wa.gov/sites/default/files/legacy/Documents/1500/SmallNumbers.pdf?uid=6425e751995 51](https://doh.wa.gov/sites/default/files/legacy/Documents/1500/SmallNumbers.pdf?uid=6425e75199551)

Confidence Interval (CI)

While currently not a requirement for publishing PRAMS data, we also highly recommend assessing, and potentially including 95% Confidence Intervals (CIs) along with your estimate. CIs are used with survey data to give an estimate of how accurately you can generalize from samples to a larger population assuming the data are not biased. Specifically, the 95% confidence interval gives the range that should contain the true population value 95% of the time. Most statistical software suitable for complex survey analyses will have an option to calculate 95% CI.

Limitations of PRAMS Data

PRAMS provides valuable population-based data on an ongoing basis for numerous uses. As with any study design, PRAMS data has some limitations, and these limitations must be considered when planning analyses and presenting results.

- *Limits to generalizing to the study population (selection bias):* May not be accounted for in the analysis weight could lead to results that are inaccurate for the population or

underrepresented subgroups. Additionally, linguistic minorities may be underrepresented since the survey is only available in English and Spanish.

- *Inability to generalize to related populations*: It is important to remember that PRAMS only includes in-state residents who had a live, in-state birth. Findings from PRAMS cannot be generalized to other populations excluded from the sampling frame.
- *Limitations of PRAMS information (information bias)*: each data source included in PRAMS has problems with missing and potentially inaccurate data due to the nature of data collection. For example, questionnaire data are self-report and may be subject to recall bias or respondents' unwillingness to report behaviors accurately.
- *Confounding*: Many factors can affect an observed association between two variables. PRAMS includes information on many of these factors but may not include all of them.
- *Lack of Power*: Standard errors for PRAMS estimates are frequently larger than they would be from a simple random survey. Lack of power is a particular concern when an event or behavior is very rare.

Additionally, CDC states PRAMS data may not be suitable for all uses even when it contains relevant information. Following are some types of analyses that may require special considerations:

- *Etiologic analyses*: These analyses investigate a possible causal relationship between a health event or behavior and a health outcome. PRAMS data might not be collected in a way to establish a clear temporal relationship or may not collect all important confounding factors.
- *Program or Policy Evaluation*: Thorough program evaluation requires several pieces of information. PRAMS cannot usually supply information on interventions targeted to small populations, or whether an intervention was administered effectively and correctly.
- *Small area analysis*: PRAMS is designed to provide statewide estimates and may not be suitable for analyses of very small geographic areas. See section on sample size for more information.

PRAMS Data Analysis

Appropriate Software

It is important to use analytic software that can account for PRAMS' complex survey design. Recommended software include SAS, STATA, R, SUDAAN, and SPSS (Software for the Statistical Analysis of Correlated Data). Basic code can be found for these software programs below in the *Sample Code* section.

Setting up your Dataset

Prior to survey analysis you must setup your dataset to account for survey respondents, weighting, strata, and combining annual files, if applicable.

- Restrict the dataset to only those who responded to the survey
    - **INQX=1** indicates participant has sufficiently completed the survey to be counted as a respondent
        - This is already applied to Limited Dataset

- Weight and Strata variables
    - Weight variable: **wtanal**
    - Strata variable: **stratumc** (single year), **sud_nest** (multiple years)

- Subset dataset to specific year(s)
    - Variable to identify birth year: **certyear**

- Include Finite Population Correction (FPC) to identify stratum sizes and indicate that the sample was selected without replacement (participants can only be sampled once)
    - Variable to identify stratum sizes: **totcnt**
    - See below on how to use with different survey software

PRAMS Variables to Produce Common indicators

Your analysis will determine how to define specific indicators. Here is how DOH defines a few common indicators.

Most dichotomous No/Yes variables in the PRAMS survey are coded **1=No** and **2=Yes**.

Unintended Pregnancy (UIP) – PRAMS pregnancy intention variable (**pgintent**) is used to inform unintended pregnancy estimates in WA. Please note using PRAMS alone does not account for unintended pregnancies that do not result in a live birth. DOH also uses number of abortions in their UIP estimate, although not all abortions are the result of an unintended pregnancy. DOH's calculated variable for pregnancy intention is **pgint3** and is coded as:
- Intended: pgintent = 2 (wanted to be pregnant sooner) or 3 (wanted to be pregnant then)
- Unintended: pgintent = 1 (wanted to be pregnant later) or 4 (did not want to be pregnant then or any time)
- Unsure: pgintent = 5 (was not sure)

Breastfeeding - Breastfeeding duration (including pumping breast milk) is assessed at time of survey, which is at least 8 weeks postpartum. Breastfeeding can be calculated several different ways, including

ever having breastfed (**BF5ever=2**), breastfed for 8 or more weeks, or other categorization. CDC creates an analytic variables for number of weeks breastfed (**bf5weeks**), and DOH uses the following categories: 8+ weeks, 4-7 weeks, 1-3 weeks, <1 week, Didn't breastfeed.

Race and Ethnicity – DOH categorizes into seven categories – Hispanic, NH* White, NH Black, NH American Indian/Alaska Native (AIAN), NH Native Hawaiian/Other Pacific Islander (NHOPI), NH Multiracial. This categorization is provided in the dataset as **race7**, a DOH calculated variable.
- Hispanic: mom_hispanic = 1, 2, 3, 4, 5
- NH White: mother_race_summary_code = 10 (White)
- NH Black: mother_race_summary_code= 12 (American Indian/Alaska Native)
- NH Asian: mother_race_summary_code = 13 (Asian)
- NH NHOPI: mother_race_summary_code = 14 (Native Hawaiian or Pacific Islander)
- NH Multiracial: mother_race_summary_code = 20 through 44, 55 (all 2+ races reported. see Data Dictionary for details)

*NH= non-Hispanic

Sample Code

Results may be slightly different depending on software package used. The point estimates will be consistent across all packages, but the estimation of standard errors and confidence intervals might differ because each package uses a different estimation approach to estimate variance. There is no "right" or "wrong" approach and the difference across packages is fairly minor.

There are many ways to write code – this is just one option to get you started in SAS, STATA, R, SUDAAN, and SPSS and produces comparable results.

Setting Up the Analytic Dataset

SAS Complex Survey

1. Combine the single-year/single-state weighted PRAMS SAS datasets into a multiyear analysis dataset
    a. Sort the dataset by variable SUD_NEST:
    ```
    PROC SORT DATA=<pramsdata>;
    BY SUD_NEST;
    RUN;
    ```
    b. 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;
table <variable of interest> / row cl (TYPE=LOGIT);
run;
```
      a.   NOTE: When running confidence intervals (CIs), use LOGIT CIs (TYPE = LOGIT)

STATA

Svyset manages the survey analysis settings of a dataset and is used to designate variables that contain information about the survey design, such as the sampling units and weights. Use the following statement when setting up your dataset in STATA:

svyset _n [pweight=wtanal], strata(sud_nest) vce(linearized) singleunit(missing) fpc(totcnt)
set more off

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.

SUDAAN

Use the following statements when using SUDAAN modules:

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

SPSS

Use the following steps to set up the PRAMS sampling design in SPSS:

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

Assessing Response Numbers and Rates

To assess number of respondents and response rates, look at the unweighted (raw) frequencies of the sub-population(s) and variables of interest. You can do this even before you set up your analytic dataset. Here is basic code to produce unweighted frequency counts in SAS and STATA.

SAS

Frequency of one variable:

```
proc freq data=<pramsdata>;
table <variable of interest> / missing;
run;
```

Frequency by demographic or other variable of interest:

```
proc freq data=<pramsdata>;
table <variable of interest> * <variable of interest> / missing;
run;
```

STATA

Frequency of one variable:

```
tab <variable of interest>, missing
```

Frequency by demographic or other variable of interest:

```
tab <variable of interest> <variable of interest>, missing
```